The development of extraterritorial human rights safeguards as a strategic tool in foreign policy-making – American, German, and British approaches towards the international prohibition of torture since 9/11

Cumulative thesis submitted for the degree of "Doctor rerum politicarum" in Political Sciences (Dr. rer.pol.)
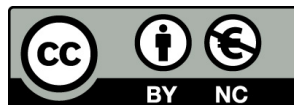
Faculty of Social Sciences, Economics, and Business Administration

University of Bamberg

Janina Heaphy, M.A.

Bamberg, 2023

The development of extraterritorial human rights safeguards as a strategic tool in foreign policy-making – American, German, and British approaches towards the international prohibition of torture since 9/11

Cumulative thesis submitted for the degree of "Doctor rerum politicarum" in Political Sciences (Dr. rer.pol.)

Faculty of Social Sciences, Economics, and Business Administration
University of Bamberg

Janina Heaphy, M.A.

Bamberg, 02/06/2022

Members of the Doctoral Committee

1) Prof. Dr. Monika Heupel
   Otto-Friedrich-Universität Bamberg

2) Prof. Dr. Thomas Saalfeld
   Otto-Friedrich-Universität Bamberg

3) Prof. Dr. Ruth J. Blakeley
   University of Sheffield

Date of Defense: February 10, 2023

For Grandad and Chuck,

'til we meet again.

Table of Content

Framework Paper

0. Formal Requirements

The submitted dissertation is in fulfillment of the university's requirements for receiving the degree of a "Doctor rerum politicarum" (Dr. rer. pol.) as outlined in the following guidelines:

- Promotionsordnung der Fakultät Sozial- und Wirtschaftswissenschaften der Otto-Friedrich-Universität Bamberg vom 31. März 2008

- Leitlinien für kumulative Dissertationen des Promotionsausschusses der Fakultät Sozial und- und Wirtschaftswissenschaften, Stand Februar 2013

- Handreichung der Fachgruppe Politikwissenschaft zur Anwendung der Promotionsordnung bezüglich kumulativer Promotionen im Fach Politikwissenschaft in der Fassung von Mai 2019

The dissertation consists of three individual research papers and one comprehensive framework paper of 43 pages. Each article has been subject to peer-review processes in SSCI-ranked journals; two of the three have been accepted for publication whereas the third paper is currently still under review. While one article has been co-authored, the other two articles have been single-authored fulfilling thus all requirements for a paper-based dissertation as stated in the guidelines mentioned above.

The first paper, "Seeing Reason or Seeing Costs? The United States, Counterterrorism, and the Human Rights of Foreigners" has been published in *the European Journal of International Relations*, which is officially SSCI-ranked and bases its work on peer-review processes (5-year Impact Factor: 4,402). The second paper, "When Identity Meets Strategy – The development of British and German anti-torture policies since 9/11" has been submitted to *International Studies Perspectives* (5-year Impact Factor 2,500) and is currently under review. Finally, the third article "British Counterterrorism, the international prohibition of torture, and the Multiple Streams Framework" has been published in the SSCI-ranked and peer-reviewed journal *Policy& Politics* (5-year Impact Factor: 3,875).

In addition to the second and third article, which have been single-authored, the first article was co-authored with Monika Heupel and Caiden Heaphy, both from the University of Bamberg. While the development, conceptualization, and writing of this first article were done together as a team, the research for two of the paper's three empirical case studies, namely on torture and targeted-killing, were the explicit focus of the defendant of this dissertation.

The following framework paper presents in detail the broader research program behind the three independently written articles, providing an overview over the studies' main results as well as the findings' broader implications.

1. Introduction

Why do powerful states introduce human right protections to foreigners abroad and thus commit to principles that they have previously repeatedly violated in the past? Since 2001, the United States of America, the United Kingdom, and Germany, alongside various other states, determinedly engaged in a fight against terrorism, in order to not only bring Al-Qaeda and the Taliban to justice, but to also prevent any attacks similar to 9/11 from happening again. Yet, while continuously condemning the terrorist groups' behavior, the states' own conduct often stood in a stark contrast to their obligations under humanitarian law and the international human rights regime: Breaches of the right to privacy, the right to be free from torture, and the right to life are only a few examples of the numerous human rights violations committed during Western counterterrorism operations. While the harm of such policies, in particular the infamous US Central Intelligence Agency's (CIA) Detention and Interrogation Program, has received considerable attention, little research has analyzed the reasoning behind the eventual development of corresponding protection provisions for afflicted foreigners abroad. In exemplifying respective processes by scrutinizing the evolution of American, British, and German commitments to the Convention against Torture (CAT) and by offering some complementary insights into similar US safeguards against mass surveillance and targeted killings, this dissertation provides a multi-dimensional analysis of the drive and motivation behind each state's respective counterterrorist policy-making. By scrutinizing why such safeguards form in different national contexts, this thesis provides further insight into how Western countries, despite their relative power and influence, can still be held accountable for their violations of the human rights of those "beyond" their official jurisdiction.

When looking at American, British, and German infringements on the right to be free from torture, exact numbers and comprehensive accounts regarding the CIA Detention and Interrogation Program and the involvement of allied states are scarce. Nonetheless, parliamentary reports and investigations by the European Union (EU) grant detailed insights into each states' various violations of the CAT. While the US, with its use of waterboarding, sleep deprivation, "walling", beatings, rectal rehydration without medical necessity, sexual assault, and mock executions (Office of Legal Counsel, 2002: 2-4; Senate Select Committee on

Intelligence, 2014: 11) mostly acted as a direct perpetrator, Germany and the UK assumed primarily assisting roles. In particular, German and British intelligence services enabled, among others, rendition flights via their respective air spaces, provided vital information for the capture of suspects, and prepared question sets for the ensuing interrogations (European Parliament, 2006).

While there were considerable efforts to keep CIA interrogations confidential, details about the torturous techniques and the states' involvement eventually trickled down to the public sparking multiple waves of outrage, but ultimately failed to translate into consistent public pressure. Starting in 2004, the Abu Ghraib Scandal with explicit pictures of detainee abuse shook the US, causing major uproar among the American people; four years later, in the presidential race between Obama and McCain, the topic of Guantánamo and coercive interrogations remained of high relevance, before eventually, the public interest into extraterritorial torture faded. In the UK, findings about the country's role in the CIA program likewise sparked severe discontent among the British people, but public mood quickly turned because of the state's rather indirect involvement and the government's strategy of denial and guilt aversion (Blakely and Raphael, 2020), branding any further allegation or actual evidence as an "insult against our brave soldiers" (May, 2016). Similarly, details about Germany's involvement gradually unfolded, but despite the media coverage and a parliamentary investigation, no major public pressure crystallized. Instead, the Merkel administration invoked an even stronger narrative of denial (Steinmeier in Deutscher Bundestag, 2009: 247), so that the topic of German assistance to the CIA mostly disappeared from the public debate.

Interestingly, however, despite the limited public interest, the ambiguous provisions in international law[1], and the continuously high terrorist threat, the US and the UK eventually decided to, at least officially, change their policies, whereas Germany, conversely, opted for policy continuance. While the US Congress and the Obama Administration bound the US military and the CIA to the Army Field Manual, which prohibited any use of torture or inhumane, cruel, or degrading treatment over the course of three individual policy introductions, the UK issued two policies to provide clearer guidance to interrogators operating with foreigners abroad. Indeed, neither the American nor British policies provide a panacea against torture, in fact, most of the safeguards come with major loopholes. Nonetheless, the states' mere recognition of the need for new safeguards is an interesting development in itself, as it

---

[1] The extend of the CAT's extraterritorial application during war, meaning whether or not the treaty obligations extend to foreigners outside the respective signatory's own territory, is still a strongly debated topic; for further details on the debate see: Gibney and Skogly, 2012; Kalin, 2007; Nowak, 2010.

ultimately reveals the actors' sudden quest to at least appear norm compliant. While traditional human rights literature can explain such a phenomenon for smaller states via the principles of international coercion and persuasion, there is little research on why more powerful, perpetrating states deliberately impose self-restrictions on their national security strategy even if only to appear norm compliant. This becomes, however, particularly interesting when considering the countries' different strategies of policy-change, as well as the contrasting example of the German Bundestag deciding to maintain the same rules and laws in place that enabled the CAT violations in first place.

Through the example of the numerous CAT violations and the states' subsequently differing policy trajectories, this thesis seeks to answer the general research question of why powerful states opt to restrict their counterterrorist operations by providing human rights protections to foreigners abroad. By choosing the US, the UK, and Germany the thesis explores this question from different angles while relying on data from three Western allies that, according to their own parliaments, breached the principles of the CAT in multiple ways, despite being signatory parties. In doing so, the case selection enables a nuanced analysis on how the states' differing degrees of violations and varying national interests have impacted the final policy outcomes, shedding thus light on the following broader research themes:

- How a perpetrating state's power on the international stage affects its vulnerability to accountability campaigns.
- The role of the perceived severity of a state's norm violations in holding it accountable for its norm-violating behavior.
- Whether human rights norms are relevant on an extraterritorial level given the widespread violations against foreigners abroad.

To address these three research themes, this dissertation presents a rationalist perspective, which argues that the states based their decision-making on strategic calculations, weighing the anticipated costs of restricting their counterterrorist operations against the costs of policy-continuance. In this context, the research findings stress the significant impact that the anticipation of future costs has on policy-making. Accordingly, decision-makers did not merely focus on immediate or already materialized consequences, but rather based their policy-choices on predictions made by experts from the respective parliaments, intelligence communities, or militaries. Hence, the most promising way of holding powerful states accountable lays within manipulating their cost-benefit analyses; either by engaging in targeted consultations or by presenting arguments, which credibly challenge the long-term success of their carefully crafted

legitimacy and cooperation claims on the international stage. In sum, the different policy outcomes can be explained by looking at how the state's choices enhanced their respective strategic interests, and how, in an attempt of holding the countries accountable, internal and external actors have been more or less successful in manipulating such strategic calculations.

Relatedly, the dissertation maintains that depending on the position and power a state holds on the international stage, mechanisms of social influence in traditional accountability campaigns, like shaming or strategic learning, unfold differently. In the context of this study, the aforementioned mechanisms are understood as processes of intervention where internal or external actors attempt to influence the decision-makers with either (a) leaked information that tarnishes a state's reputation or (b) with new strategic arguments that try to manipulate the decision-makers' calculations and to thus trigger norm-abiding learning processes. Consequently, unlike most existing human rights literature focusing on transitioning states, this dissertation claims that it is not preferential trade agreements or foreign investments that can convince backsliding powerful states of restoring their treaty commitment, but rather strategic arguments regarding efficiency and legitimacy. While the first point is likewise applicable to transitioning states, the second point requires further differentiation. In contrast to transitioning states who usually gain legitimacy immediately by committing to a new treaty, a similar policy-change for most backsliding powerful states often first requires an admission of guilt before then eventually translating into legitimacy gains. Correspondingly, accountability campaigns must take such notions into consideration, especially as the targeted states usually justify their elevated position and actions with a carefully crafted image of being international law-abiding state actors. Hence, if wanting to trigger learning processes, it is this image's longevity, which intervening actors have to credibly put into question by showing how the anticipated legitimacy losses outweigh any costs stemming from self-determined guilt admissions. Similarly, powerful states' also exhibit a greater ability to circumvent direct shaming campaigns as they cannot only evoke narratives of denial (see Blakely and Raphael, 2020) or stigma management (Adler-Nissen, 2014), as transitioning states can, but they can also use their power to force other state actors to echo such notions, and to thus not only gain rhetorical supporters, but to also reach a larger audience.

To a similar extent, and unlike strategic learning processes, shaming campaigns are impacted by the perpetrating states' perceived severity and (in-)direct degree of norm violations; due to which the dissertation argues that in case of indirect violations shaming campaigns should rather undermine a state's reputation of being a reliable cooperation partner than solely its public image of being a human rights abiding actor. According to this study, strategic learning

campaigns have to be tailored to the targeted states preferences in any case, so whether violations took place in a direct or indirect manner might affect the preparation and framing of the campaign, but should not necessarily hamper respective efforts as strategic campaigns that target indirect violations can also create anticipatory notions of future strategic losses. Regarding shaming campaigns, however, existing literature (see e.g. Keck and Sikkink, 1998: 205) generally agrees that graphic and explicit evidence is needed to trigger compelling public pressure; in case of indirect violations, however, such evidence is usually hard to produce diminishing thus the chance of public uproar. Therefore, the dissertation suggests that instead of solely focusing on the targeted state's public, respective campaigns would benefit from also targeting the state's reputation among its allies and other states as ultimately, all states are dependent on well-functioning security cooperation, and thus, a decrease in reliability and mutual trust would come at a great cost for the indirectly violating state.

Finally, the dissertation also shows how normative persuasion as a tool of intervention and social influence has failed, but that nonetheless, human rights norms impacted the states' policy trajectories in three important ways. Firstly, while norms, beliefs, and values might not have convinced the respective decision-makers of policy-change on their own, they nonetheless might have motivated the intervening actors to pursue their campaigns and to lobby for protection policies that benefit foreigners abroad. Secondly, understanding country-specific norms, values, and identities enables the intervening actors to better tailor their strategic argumentation to the individual state preferences, strengthening thus not only their line of reasoning, but also allowing for a better basis for subsequent learning processes. Ultimately, the dissertation also demonstrates that norms matter in the perception of costs, by illustrating how the scrutinized states considered future reputational losses as prohibitive factors, rather than administrative costs, such as the financial spendings necessary to draft and pass a new bill.

As a result, the study, while remaining relevant for policymakers and for the broader public, also contributes to various academic debates. Although human rights literature offers an extensive debate on norm-compliance and commitment (Allendoerfer et al., 2020; Bassano, 2014; Risse et al., 2013; Sanders, 2018; Shor, 2008; Simmons, 2013), it ultimately falls short in providing a comprehensive explanation for extraterritorial commitments. Likewise, prominent literature on intelligence studies and counterterrorism links a state's treaty commitment to their individual counterterrorism paradigms, leaving aside, however, any particular focus on the motivation behind paradigm-changes, as well as any extraterritorial components they may have (Andreeva, 2021; Brimbal et al., 2019). Torture specific research, in turn, predominantly focuses on how torture policies got introduced into the 'War on Terror'

(Arsenault, 2017) and on how various states tried to avoid or downplay respective allegations in the aftermaths (Blakely and Raphael, 2020), a comprehensive outline encompassing the later policy developments within multiple different states, however, is hitherto missing. Despite the little attention that the various states' turn towards extraterritorial safeguards has gotten so far in the academic debate, the topic is still highly relevant. Only a solid understanding of the states' motivation behind such actions can enable other (non-)state actors to hold powerful states accountable, and to thus enforce the international ban on torture.

Additionally, the dissertation presents three central theoretical contributions relating to the importance of anticipatory notions, the role of foreign policy interests in domestic policy-making processes, and to the debate on strategic versus norm-based state behavior. Firstly, the dissertation stresses the value of using anticipated, country-specific costs in argumentative interventions, as the presentation of imminent costs frequently triggers decision-makers to display defensive behavior, rather than giving them time to process the presented arguments and to thus facilitate strategic learning and incremental policy-change. In addition, by using a theory-building approach, decision-makers' foreign-policy interests are being inserted into traditional models for domestic policy-making. In doing so, the dissertation claims that a politician's sought for power is not only restricted to the domestic level, but that in case of a credible threat to their country's and personal standing on the international stage, decision-makers will even overrule the opinion of the domestic public if it ensures the preservation of their international negotiation power. Finally, the dissertation also contributes to the theoretical debate (see e.g. Hafner-Burton and Tsutsui, 2005; Risse et al., 1999) on whether states act out of strategic calculations or norm-based beliefs by outlining how all three states followed the presumed gain-maximizing strategy, without negating, however, how norms played an important role in shaping the states' strategic preferences.

Structure of the Dissertation

Given the gap in literature, the dissertation seeks to answer the overall research question of why formerly non-compliant, powerful states do - or do not - eventually introduce protection policies for foreigners abroad. In doing so, the work particularly analyzes the decision-making processes preceding policy-change or policy-continuance, by scrutinizing the motivation of state actors, and exploring under which circumstances Western states are more or less likely to alter their policies for the benefit of non-citizens abroad, as outlined in Figure 1. According to this framework, the main dependent variable in each study is the states' final decision in favor or

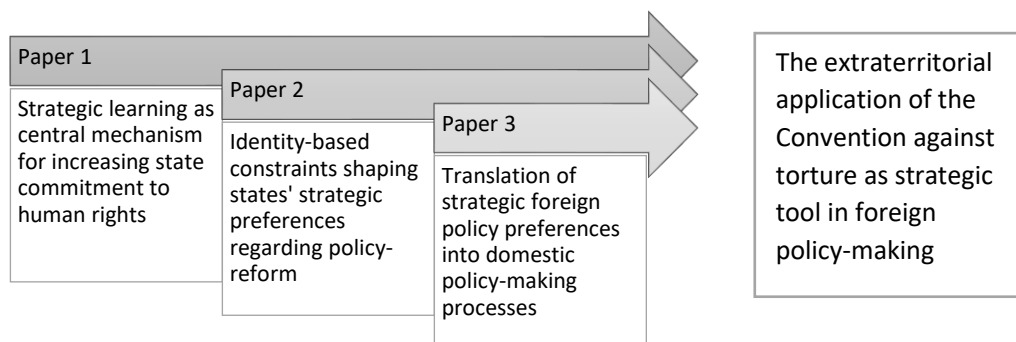against policy reform. Each paper, however, provides different insights into this leading research question.

Figure 1: Overview and structure of the dissertation.

The first paper scrutinizes how the US had been held accountable for its CAT violations during the 'War on Terror', while also offering additional insights into similar processes regarding the right to privacy and the right to life. In particular, the paper traces three deductively derived causal mechanisms of social influence, namely normative persuasion, coercion, and strategic learning, throughout the development of US protection policies for foreigners abroad; testing thus, their impact and applicability. In doing so, the article concludes that in the context of big scandals, the US can indeed be coerced into policy-change. Outside such crises, however, an internal or external actor's argumentative intervention which triggers strategic learning processes among the decision-makers is most promising in achieving change. Normative persuasion did not seem applicable in any of the cases.

Subsequently, by enrichening the overarching rationalist postulations with insights from identity literature, paper 2 zooms in on how states, based on country-specific, identity-based constraints, develop their strategic preferences, and how they decide whether the introduction of extraterritorial safeguards is in their best interest. Accordingly, paper 2 focuses on Germany and the UK as significant parallels between potential explanatory factors exist in both cases, and yet, both countries ultimately embarked on opposite policy trajectories. Hence, for the period between 2001 and 2010, the article scrutinizes each state's decision-making for national, international, and elite constraints to analyze why Cameron saw a reform of the British interrogation guidelines as beneficial, while Merkel rather opted for denial and policy-continuance. The study shows that especially patterns within history, culpability, and international self-understanding played a decisive factor in the diverging development of British and German strategic preferences.

In conclusion, Paper 3 focuses on the latest of the British policy reforms from 2019, exploring why Theresa May, in the absence of electoral pressure, scandals, or ideological concerns, saw the strategic need to alter existing interrogation guidelines. By choosing this single case study, this article enables an in-depth analysis of micro policy-making processes among British policymakers. Specifically, this paper illustrates how a politician's pursuit for international negotiation power affects domestic decision-making procedures, and how this in turn, can be manipulated in favor of extraterritorial policy reforms. Although such interests are not necessarily indicative of norm-compliant behavior, findings from this paper provide insight into the potential accountability efforts, even in an otherwise unfavorable context for traditional policy reform.

In sum, this dissertation contributes to current literature by filling a theoretical as well as empirical gap. The thorough testing and analysis of applicable mechanisms of social influence within the context of the US offers answers on how to hold powerful states accountable for extraterritorial human rights safeguards. Subsequently, the dissertation offers a theoretical approach on how to incorporate these extraterritorial notions into existing policy-making theory; a field that so far predominantly focuses on domestic policy influences with little regard to foreign policy interests. By including the UK and Germany into the analysis, the dissertation likewise offers additional comprehensive empirical insights to the topic of torture and complicity in the post 9/11 era; a topic, which especially in the case of Germany has so far been largely neglected. Furthermore, the mix of cases with direct and indirect violations as well as differing policy outputs allows for a more holistic picture and better understanding of the states' final strategic maneuvering.

The remainder of the framework paper is structured as follows. Section 2 systematically reviews relevant literature on states' human rights commitments, counterterrorism, and torture in order to situate the dissertation into a broader research program and to identify relevant research gaps. Section 3 presents the overall theoretical framework providing further information about strategic decision-making processes in politics and summarizing respective key concepts and assumptions. Section 4 presents the general research design, before Section 5 delves into the individual articles and highlights the respective findings and overarching research results.

2. *Literature Review*

Since the early work of Donnelly (1989) and Forsythe (1991) the study of human rights forms an integral part of the field of International Relations. Although the topic has drawn a lot of

scholarly attention over the years, crucial gaps in literature remain, especially with regards to the introduction of human right safeguards for foreigners abroad. Suitably, research on human rights offers a broad spectrum of explanations on why states even violate the basic rights of their citizens (e.g. Davenport, 2000; Poe, 2004); generally concluding that in one way or another the fear of losing military or political power, triggers the respective officials to repress their opposition (Des Forger, 1999; Straus 2007). While domestic human rights violations still play an important role in new strands of the literature (Hollyer and Rosendorff, 2011; Neumayer, 2005; Vreeland, 2008), most scholarly work turned towards international actors and the ever-evolving understanding of human rights (e.g. Hafner-Burton and Tsutsui, 2005; Forsythe and McMahon, 2016; Peksen, 2012; Piccone, 2018). Under the primacy of Constructivist and normative approaches (see Bates, 2014; Goodman and Jinks, 2013; Murdie and Davis, 2012; Risse et al., 1999), scholars have focused on the ripple effects of human rights violations committed and enabled by (non-)state actors, transnational actors, or by international organizations, showcasing ultimately how, especially during times of conflicts, human rights violations are not always tied to territorial boundaries nor to the sole agency of states (see Clapham, 2006; Heupel et al., 2018; Karp, 2014).

Hence, the nexus of counterterrorism and extraterritorial human rights violations featured in this thesis mirrors key patterns of the current academic debate. Despite the depth of this field of research, however, relevant compliance and commitment theories fall short in explaining why hitherto backsliding states (i.e. states that have already been considered compliant signatories to a treaty, before then displaying non-compliant behavior) suddenly move towards introducing protection policies for foreigners abroad. In the context of treaty commitment, opinions differ sharply with one group of scholars emphasizing the importance of norms and socialization, and another group reiterating rationalist assumptions (Cardenas, 2004). The latter approach entails the classical enforcement dilemma, claiming that states "talk the talk" (i.e. signing human rights treaties without intention of committing), but that without an authority controlling and enforcing these rules, commitments only serve the purpose of "window-dressing" (Hafner-Burton and Tsuitsui, 2005: 1378). Accordingly, scholars claim that states solely act upon their own interest (Goldsmith and Posner, 2005), committing hence only to treaties if they anticipate strategic gains or it enables the state to avoid future costs (Cole, 2009). Corresponding research focuses explicitly on states' initial decision on whether to sign a treaty, whereas the strategic gains of restoring commitment to an already ratified treaty, in particular in the context of extraterritorial obligations, remains largely unexplored.

The other approach towards human rights commitment, featured within Risse et al.'s (1999) Spiral Model or Keck and Sikkink's Boomerang Model (1999) for instance, emphasizes the values of norms. In doing so, it particularly relies on the input of transnational actors, who from inside the perpetrating state transfer information about violating behavior to actors abroad, who then in turn, via shaming, persuasion, or bargaining, pressure the perpetrator into committing to international protection standards (Simmons, 2013). The Spiral Model further traces how the state gets eventually entrapped in its treaty commitments internalizing eventually respective human rights norms as a general liberalization process unfolds in the perpetrating country (Risse et al., 1999). While both models account for the possibility of internalization processes being interrupted and terminated prematurely, they lack, with the exception of Sikkink (2013), specification on backsliding behavior by states who had already completed all five steps of the spiral successfully. As a result, recent scholarly work focuses particularly on states who have already created protection mechanisms for their citizens, but who ultimately returned to violating the very same safeguards they had previously created (Allendoerfer et al., 2020; Bassano, 2014; Risse et al., 2013; Shor, 2008). In spite of this, the extraterritorial notion of compliance and commitment is also barely considered in this strand of research.

In this context, however, it is likewise key to distinguish whether the impetus for rule-conformity derives from internal or external stimuli, making a solid understanding for the international community's ability to hold perpetrating states accountable key. The first central accountability mechanism discussed in the academic debate pertains to the method of sanctioning; by invoking financial sanctions, for instance, international organizations or individual states can manipulate the targeted state's dependence on goods, market access, and capital flow in order to force the perpetrating actor back to norm-compliant behavior (Donno and Neureiter, 2018; Hafner-Burton, 2005). Similarly, using shaming methods to trigger public outrage or to target a state's reputation is a frequent tool for holding a state accountable. By laying norm-deviant behavior bare, the shaming actor forces the targeted state to publicly acknowledge their misconduct, jeopardizing thus directly the state's international credibility and cooperation (Hafner-Burton, 2008; Krain, 2012; Murdie and Davis, 2012). In the end, however, less drastic methods can also contribute to externally increasing a state's commitment to international norms, as persuasion and bargaining, for instance, can either invoke normative pressures or establish clear bottom lines during negotiations (Finnemore and Sikkink, 1998; Reinold and Zürn, 2014).

When looking at the above-discussed literature on state commitment and accountability, it becomes apparent, however, that despite in-depth and diverse research over the years, a

significant gap remains. With the exception of Heupel's (2020) work on indirect accountability, existing theoretical models still focus predominantly on states violating their own citizens' rights, excluding thus extraterritorial violations and falling consequently short in tracing respective commitment processes. Likewise, only few authors applied commitment or accountability models to powerful states, as opposed to transitioning states, which is however important as the preconditioned dependencies and vulnerabilities differ greatly between those two sets of states. Even in Sikkink's work (in Risse et al., 2013), which does indeed test the Spiral Model in the case of US torture during the Bush Administration, there is no clear indication of motivation behind the American policy-change; namely, for why the US eventually decided to implement new safeguards. Finally, different levels of commitment have been tested on a great variety of countries (e.g. Cardenas, 2007), but when it comes to indirect perpetrators like the UK or Germany, there are barely any in-depth studies of the states' policy trajectories (for an exception see Blakely and Raphael, 2017), even though such studies could contribute to a better understanding for the states' decision-making.

Even when shifting the focus from human rights literature towards the field of intelligence studies, similar patterns and shortcomings can be observed. In general, most scholarly work within the field of terrorism, does agree that the consistently elevated terrorist threat and the continuous spread of terrorist networks throughout Asia, the Middle East, and Africa, are ultimately a testimony to the incomprehensive, and ineffective nature of most Western states' deterrent counterterrorist strategies (Davies, 2018; Mir, 2018; Zulaika, 2009). Accordingly, the lack of pre-emptive and moral approaches towards fighting terrorism does undermine the states' capabilities and efforts, leading thus to scholars calling for a change of strategy emphasizing, in particular, the urgency of taking ethical standards into consideration (Clifford, 2017; Coates, 2016). In doing so, the respective literature clearly reiterates why an overhaul of counterterrorist-related policies is needed, but does not provide comprehensive nor detailed information on how such a policy-change can be prompted, or administered.

Additional work within the field of intelligence studies and torture, however, complements these notions of change by delving into so-called paradigm changes, whose exploration provides a general understanding for why countries might shift their counterterrorist strategies, but ultimately still leaves empirical gaps. Suitably, large parts of the relevant literature analyze how 9/11 and the attack's aftermaths affected the US' approach towards terrorism. By dissecting Bush's and other senior official's speeches, scholars like Jackson (2005, 2015), or Jacobs (2017) for instance, show, how the politicians' reactive as well as security-centered rhetoric quickly transpired into tangible policy alterations pertaining to the military's and the

CIA's interrogation practices abroad. This focus on critical junctures and crises spans like a red thread throughout literature explaining the changing nature of counterterrorist paradigms (Andreeva, 2021: 755; Arsenault, 2017). Hence, existing work can comprehensively explain why after major terrorist attacks states suddenly begin to condone coercive interrogation techniques (Luban, 2007), but with one exception (Blakely and Raphael 2017) the field lacks detailed accounts of intelligence-related paradigm changes outside of crisis modes or that can explain why states like the UK or the US decided to at least partially recognize extraterritorial treaty obligations. While Blakely and Raphael (2017) do indeed investigate the UK's motivation behind their spurious policy-change, but only provide important insights regarding a single case study; and therefore lack contrasting inferences on the policy-choices of other powerful states.

In sum, the thesis is well embedded within academic literature reflecting themes and trends within current debate, while offering new insights on Western states' approaches towards extraterritorial human rights. In doing so, the thesis contributes in three ways to current debates: First, by focusing on the development of American, British, and German extraterritorial human rights safeguards, the thesis enriches debates on states' human rights commitments by looking at the countries' behavior beyond their own border, while also contrasting positive and negative cases of policy-change. In terms of accountability literature, some of the mechanisms discussed above (i.e. facets of sanctioning, shaming, normative persuasion, and bargaining) are being tested for their effectiveness against powerful states and in cases of indirect violations. Lastly the general accumulation of inferences from across all case studies allows for a better understanding of multiple countries' gradual paradigm changes outside of times of crises.

3. *Theoretical Framework*

The following section presents the general theoretical framework of this dissertation, assuming in particular a rationalist approach of commitment to understand the states' decisions on whether to provide protection safeguards foreigners abroad, which despite the individual papers' different theoretical approaches in the individual papers spans like a red thread throughout all three articles and the states' decisions on whether to provide protection safeguards foreigners abroad. By following a rather strategist than normative view, the thesis primarily examines the states' decision-making, while recognizing the new policies' limited scopes and loopholes. Although these policy shortcomings are evidence of the lack of current constitutive effects and socialization processes related to extraterritorial human rights, the thesis

does not inherently reject the importance of norms. Rather, norms still play a central role in motivating individual policy entrepreneurs, establishing regulative effects on the international stage, or informing the political culture behind a state's decision-making (see Duffield, 1998). Hence, to build a comprehensive understanding for the overarching theoretical approach and the respective unit of analysis, the subsequent section first elaborates on the nature and key characteristics of the dissertation's central actors, before then offering further details about the actors' strategic preferences and the basic elements of strategic decision-making processes.

In line with rationalist strands in foreign policy-making, the dissertation's unit of analysis and thus the principal actors scrutinized in the thesis are state leaders or other high-ranked decision-makers within a country that hold the political power and authority to introduce new policies and to represent the state on the international stage (Alden, 2017). As such, the actors are assumed to act rationally, meaning that they instrumentally engage in politics based on the information they have available as well as their national interests and preferences, to pursue the most cost-effective means to reach their goals (Fearon, 1995; Simon, 1997). In doing so, they display risk-adverse and evaluative behavior, and are capable of taking informed decisions and of seeking equally beneficial alternatives, if necessary (Epstein, 2013: 293). Particularly salient in this context is the mostly self-interested approach towards decision-making, in which powerful actors ultimately aim to secure the greatest benefits for themselves. Nonetheless, patterns of altruistic, normative or ideational influences can also be traced within the rational actor's conduct as long as they do not directly contradict his or her own goals (Snidal, 2002: 87).

Considering the principles of bounded rationality and thus the limitations of perfect rational decision-making, it is important to differentiate between externally given limitations and the rational actors' intrinsic preferences, meaning the hierarchy of favored outcomes of an interaction (Frieden, 1999: 39, Shannon et al., 2019; Simon, 1997). Accordingly, depending on the other actors involved, the information available, the individual preferences voiced, and the distinct expectations held, the environment in which decisions are taken and how such interactions unfold can impose external constraints on the decision-maker's options (Hafner-Burton et al., 2017: 19). In contrast, each actor develops intrinsic strategic preferences based on national and individual interests, which do not only reflect the perceived ideal result of each decision-making process, but also a ranking of less favorable options, which are decreasingly compatible with the actor's interest (Scharpf, 2018:79). In the end, each actor strives, under the caveats of bounded rationality, for maximizing the utility of their interactions with others. Due to the interweavement of decision-making procedures and a constantly evolving strategic

environment, however, an actor's strategic preferences are not fixed, but can change over time (Scharpf, 2018: 64). Furthermore, actors can learn from previous experiences adapting their expectations and preferences accordingly (Snidal, 2002: 88).

When looking at the actual decision-making processes, the actors' preferences play a central role as they ultimately frame the anticipated costs and benefits of a policy decision. Hence, based on the individual order of preferences, the availability of resources, and within the constraints of the strategic environment, each actor analyzes the anticipated costs and benefits of the pending decision, seeking to maximize the utility of their actions in order to gain the greatest profits possible (Alden, 2017). In this context, costs are defined as any policy outcome contradicting or hampering the state's key interests, whereas benefits constitute policy outcomes that advance the state's strategic goals. As illustrated in the section on state preference, how a state defines a 'cost' or a 'benefit' is dependent on a number of key contextual factor and is subject to change over time. Within the context of national security interests, for example, the implementation of reforms governing interrogation techniques may be on the one hand perceived as an additional cost in that it reduces access to key intelligence for counterterrorism operations; and on the other hand as a benefit in that the absence of such reforms may provide fuel for negative terrorist propaganda. Which definition a state decides to prioritize depends on the capacity and the willingness of policy actors to promote respective perspectives and to provide corresponding information.

Although, most decisions especially in the context of international politics are inherently "depend[ent] on the choices of others" (Snidal, 2002: 88), the degree of such dependence is usually dictated by the hierarchy and power distribution within the respective political system. In its most traditional form, power is defined as the ability to force another actor to do something they otherwise would not do (Dahl, 1957: 202). Such dynamics are particularly reflected in decision-making processes where either weaker states are compelled not to pursue their most utility maximizing strategy, or in processes where powerful states can persistently enforce their preferred outcome as any opposition by others would be immediately sanctioned (Baldwin, 2016: 170). Nonetheless, power distributions are not permanently fixed, but can indeed fluctuate over time, especially as power alone can transpire in various shapes ranging from the traditional material power to different nuances of soft power (Lake, 2011: 45-63).

In any case, forcing powerful states into cost acceptance or policy-changes that do not reflect their favored outcome, but benefit others, commonly requires joint efforts by multiple weaker state and/or non-state actors, and often relies on either enforcement measures or on incentivizing

strategies. Hence, traditional enforcement techniques seek to artificially increase the costs of the targeted actor's favorite policy option to such an extent that the initially unappealing policy-change becomes a more attractive option (Elliot, 2018). In practice, this often takes the form of material sanctions so that the financial costs of policy-continuance are being inflated so significantly that they eventually surpass the costs of policy-change. In the context of powerful states, however, this requires the existence of strong dependencies as otherwise the targeted state can circumvent the cost increase and punish the enforcing actors (Donno and Neureiter, 2018; Hafner-Burton, 2005). In contrast, incentivizing strategies rely on manipulating the benefit side of the violating state's calculations by offering new information, aid, or for example support in pending multilateral negotiations. Similarly, however, it is usually the powerful states who can best offer special incentives that offer credible enticements, which are unique to their position and cannot be replaced by other means (Risse et al., 2013; Schimmelfennig, 2005: 829).

In conclusion, the thesis assumes a broad rationalist view on the states' decision-making, while appreciating the potential for normative influences in the analysis of strategic preferences and behavior. Accordingly, each paper focuses on different aspects and features of the overarching rationalist approach (see Fig.2), with two articles also featuring some normative notions in either the tested mechanisms of social influence (see *normative persuasion* in article 1) or in the deduction of strategic constraints (article 2). Hence, the dissertation assumes that political decision-making is based on strategic contemplations, but ultimately agrees with Jupille et al. (2003) that the consideration of normative postulations can warrant additional insights into the formation and evolution of strategic preferences and perceptions.

In sum, paper 1 embeds Rationalism in one out of three causal mechanisms (*normative persuasion, coercion, strategic learning*) by tracing if and how argumentative interventions by internal or external actors can change US policymakers' minds and eventually trigger strategic learning processes. In exploring such attempts of logic reasoning, the first paper shares its focus on the evolution of strategic preferences with article 2, which in turn, uses Rationalism and notions of identity to explore the UK's and Germany's differing strategic preferences in their respective decision-making. In addition, paper 1 analyses how preferences favoring change manifest in policy-making processes, a phenomenon also explored in paper 3, which by the example of the UK shows in detail how a state's foreign policy preferences translate into domestic policy-action. In the end, however, it is this latter domestic notion of article 3 which marks a crucial difference to paper 1 and a key commonality to article 2, as only paper 2 and 3 explicitly explore the relation between policy preferences and the states' domestic environment.
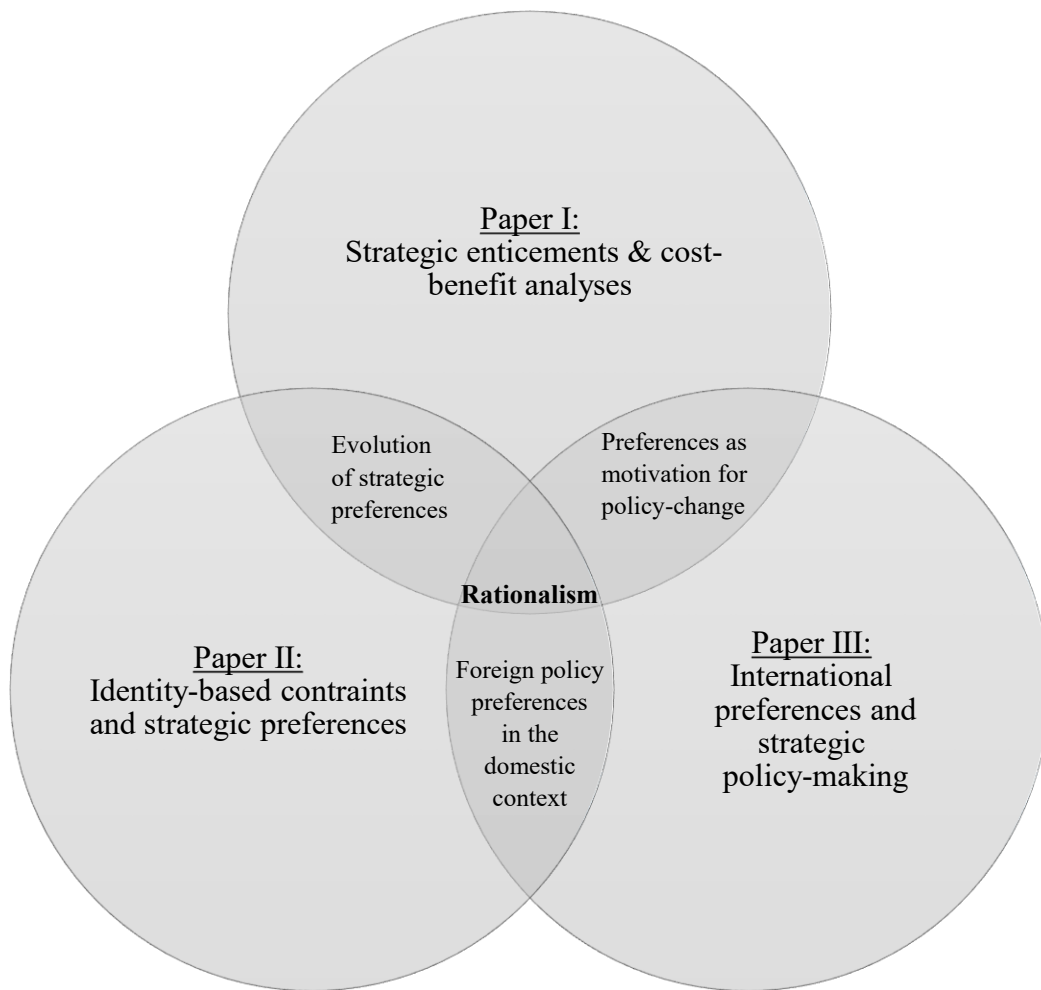
Figure 2: Rationalist concepts within the dissertation's three papers.

4. General Methodological Framework

The following section outlines the individual articles' research designs, while also discussing their basic commonalities. Regarding the latter, the dissertation as a whole consists of three qualitative papers analyzing data derived from primary sources such as parliamentary hearings, speeches, as well as from 43 semi-structured interviews conducted in Washington D.C., London, and Berlin. However, the articles' respective research designs ultimately differ from each other as deductive as well as inductive variants of process tracing and content analysis have been used to scrutinize empirical observations, and to thus answer the thesis' overarching research question. The following chapter outlines in detail each articles' units of analysis and providing further details on aspects of data collection and on the analysis of the studies' small-n designs, before then delineating how each of the three articles contributes individually to the dissertation's research agenda.

Case Study Selection

By zooming in on the 'causes-of-effects' rather than on the 'effects of causes', the research design follows a qualitative logic with small-n samples, considering also the potential impact of causal heterogeneity and equifinality (see Bennett and Elman, 2006: 457-458; Gerring, 2004: 348; Mahoney and Goertz, 2006: 229). The three papers' common unit of analysis are state leaders and other high-ranked politicians, who have the authority and political power to act and decide on behalf of the individual countries being scrutinized in the dissertation. Despite the ensuing decreased degree of generalizability, the dissertation focuses on small-n samples to maximize the use of time and means to provide thorough and in-depth inferences; this concentration of resources enables the researcher to map out complex developments in detail and to consequently identify the rationales behind the actors' decision-making (Gerring, 2004; Siggelkow, 2007).

Given the small-n approach, the guiding principle of the overall case selection is linked to how a cumulative research project can mitigate the limited generalizability of within-case inferences by comprehensively assessing multiple cases that share key features, but that nonetheless include varying contextual conditions (Beach and Pederson, 2019: 145; Bennet and George, 1997: 6). According to this approach, the dissertation centers on the US, the UK, and Germany as three powerful states which all adhere to the principles of Western Democracies; sharing thus fundamental political values and norms (Youngs, 2015: 146). The cases have been selected due to their historical commitments to multilateralism and human rights, as well as their general respect for democratic practices and institutions; hence, even though the states' characteristics differ in other aspects the shared appreciation for democratic principles and the common recognition of human rights law offer a certain degree of consistency. In line with this, each country provides avenues for domestic parliamentary investigations as well as for external actors to criticize and intervene against potential violating behaviors, both crucial pre-conditions for strategic learning and the dissertation's other mechanisms of social influence. Moreover, all three countries were identified side by side in the Dick Marty Report commissioned by the Parliament of the European Union (2006), which outlined each country's support of and involvement in CIA operations. Thus, the selection of these three powerful Western states allows for an elevated degree of consistency, especially when compared to similarly powerful autocratic states such as Russia, China, or Saudi Arabia with their different regime types, political cultures, and human rights histories.

While benefitting from consistency, however, the case selection also allows for interesting inferences based on changing variables, as the countries showcase three notably different contextual differences. Firstly, the states' degree of violating behavior varies greatly. Germany and the UK, for instance, were predominantly complicit to torturous practices, whereas the US appeared to be the main direct perpetrator whose actions have been comprehensively disclosed in leaked documents and graphic photos. In light of this variance, the nexus between blame attribution and cost-benefit analyses can be further explored as norm-driven narratives, international sanctioning regimes, and the thus anticipated costs hinge on the perception of a state's guilt and responsibility (Keck and Sikkink, 1998: 27). Similarly, the second difference alludes to the fact, that the 'War on Terror' was a conflict declared by George W. Bush, joined by the UK, and at least officially rejected by Germany. The variance in these approaches hence illustrate different ideologies and governmental security strategies, whose impact on the states' decision-making can be explored (Campbell, 1992). Finally, further contextual differences emerge when looking at the power distribution on the international level; all three states are considered to be powerful actors, and yet, the US' power to influence other states' decision-making or to withstand international pressure is higher than that of Germany or the UK (Duffield, 1998). By incorporating such variances in the case selection, the individual parameters' influence on actors' strategic maneuvering and the state representatives' decision-making can be explored in greater detail.

Furthermore, the three countries were likewise selected due to their differing policy trajectories. Whereas the US and the UK can be considered "positive cases" because they eventually introduced new policies, which at least on paper protect the human rights of foreigners abroad; Germany was evident of a "negative case" because the German government opted against such changes. The exploration of positive cases enables a very detailed analysis of the motivation and rationales behind policy-change (article 1 and 3), while also allowing for complementary comparison of respective decision-making in different policy fields (article 1). The juxtaposition of a negative and positive case (article 2), however, facilitates a better understanding for the emergence of differing strategic preferences and policy outcomes. Cumulatively, the inclusion of two "positive cases" and one "negative case" provides an opportunity to explore the potential pre-conditions for change and to scrutinize why some states perceive change as strategically beneficial, while others do not. Relatedly, the timeframe of this study spans twenty years (2001-2021) to capture all relevant policy-developments made in the respective countries.

Finally, the overarching focus on Torture as exemplary extraterritorial human right violation was the result of several conceptual and practical considerations. Firstly, the prohibition against torture is an internationally recognized, well-established human right in international treaties, of which all three states are signatories. As a result, the prohibition of torture, unlike some other human rights, is clear in its scope, and rarely disputed at its core. Despite its international recognition, however, the use of torture in counterterrorism policies by Western states remains particularly relevant, as evident by violations committed by each country and even the pro-torture statements of the Trump Administration in the US (see Trump, 2017). Secondly, the CAT violations by all three states have been extensively documented by respective UN committees, parliamentary reports, and the EU, among others. Data is therefore readily accessible to the public, and can thus be triangulated across diverse sources; thereby increasing the credibility of the research. Lastly, the selection of torture enables an assessment of direct and indirect human rights violations. Therefore, the focus on torture conveniently complements the study's broader interest in safeguards against extraterritorial human rights violations, as can tentatively be seen in Article 1.

In sum, the selection of the US, Germany, and the UK as case studies is suitable as they are relevant for addressing important empirical and theoretical gaps in the literature. First, they pose interesting cases as the countries usually portray themselves as defenders of human rights despite being perpetrators themselves. The resulting empirical findings do not only fill shortcomings in literature, but can also raise awareness of how and why Western democracies deviate from and partially return to their own principles. Secondly, despite the relatively high sensitivity of the subject, the selection of the three cases warrants access to viable information via parliamentary and media archives, as well as through in-depth interviews with policy stakeholders. Taking all aspects into consideration, the case selections offer a solid basis for answering the dissertation's guiding research questions, while taking practical, methodological, and theoretical factor into account.

Data Collection

In order to meet the standards of thorough academic research, the dissertation is based on an array of primary and secondary sources as well as on 41 interviews (see Annex 1). Accordingly, secondary sources such as academic publications have been key in gaining a general idea of the status quo and in grasping current trends in the development of extraterritorial human rights protections. Given the extraterritorial nature of the policies and the nexus to an international

conflict, a special emphasize has been put on IR literature in order to trace the various developments on the global stage; legal writings, in turn, have been particularly important in regards to understanding the (inter-)national legal framework of the countries' counterterrorism strategies. In general, a plethora of secondary information can be found on American and British violations, though only few sources deal with the nexus of German counterterrorism efforts and the infringement of the Convention against Torture. Regarding the development of the extraterritorial safeguards, some secondary sources touch on the introduction of the respective protection policies, especially in the case of the US, without thoroughly explaining, however, why two out of the three states have conceded into self-imposing restrictions on their counterterrorism strategies. Hence, the analysis of the actual intermediate steps from a CAT violation to the introduction of a safeguard still constitutes a gap in current literature.

In light of this gap, primary sources have been used to gather new information and to corroborate secondary findings. In doing so, periodic reviews by the Committee against Torture, or for instance investigations ordered by the states' own parliaments, have been used to corroborate allegations of CAT violating behavior, whereas legal documents and negotiation protocols allowed for a thorough examination of the decision-making processes scrutinized in the individual papers. Similar to the findings in secondary literature, a lot of primary material could be found in the cases of the US and the UK, whereas for Germany, primary sources mostly insinuate, for instance, a 2007 overhaul of military orders regarding the treatment of detainees abroad, but related information remains strictly confidential and inaccessible. In general, however, parliamentary investigations, debates and motions have been published by the three states, the UNCAT committees' communications are freely accessible, and respective governmental statements have been made available in archives or the internet. Moreover, inquiries and reports by the European Union provide further information, so that comprehensive access to primary sources has been warranted in all three case studies.

In addition to the analysis of primary and secondary sources, interviews with key actors and stakeholders have been crucial in gaining further information, detailed insights, and first-hand accounts regarding the country-specific motivations for policy-change or -continuance (see King et al., 2010; Leech, 2002). During the time between 2018 and 2020, a total of 41 semi-structured interviews has been gathered remotely as well as personally in Washington D.C., London, and Berlin, in order to enable a thorough analysis of the overarching research question. In addition to former senior military and CIA interrogators, John Bellinger as former legal adviser to George W. Bush, Dominic Grieve as the former General Attorney for England and Wales, as well as Guenther Nooke, the former Federal Commissioner for Human Rights Policy

and Humanitarian Aid to Angela Merkel have been interviewed. Furthermore, interviews with policymakers from the three countries' legislative and executive branches have been of particular interest due to their first-hand knowledge of the respective decision-making processes. Likewise, individuals with a significant expertise in the field, including renowned scholars from universities or think tanks, journalists, as well as NGO staffers have been asked about their experiences regarding the states' stance towards extraterritorial safeguards. Finally, members of the various security agencies and the military, who were directly involved in relevant detention and interrogation operations, like Glenn Carle or Mark Fallon, were also interviewed in order to capture their perspective on the issue and to grasp the impact of the governments' final decisions. Altogether 21 interviews have been conducted in the US, nine in Germany, and thirteen in the UK; further information about the interviewees' different backgrounds can be found in this dissertation's annex.

The interviews used throughout this research follow the ethical guidelines of qualitative research (Miller et al., 2012) so that all interviewees have been fully informed about the purpose and the object of this study; furthermore, their privacy and confidentiality remain respected. Already upon approaching potential interview partners via mail, details about the project's aim and motivation have been given, though to avoid any bias, specified information regarding the hypotheses have not been communicated until after the interviews had been finished. All interviewees have given their consent to their accounts being used for research. In order to adhere to the principles of confidentiality and to respect the interview partners' privacy, their full names will not be given in the respective citations, instead the interviews will be numbered and solely the profession as well as the location and the date of the interview will be indicated.

Data Analysis

In order to answer the overarching research question adequately, each paper follows a different qualitative methodological approach so that in accumulation a cohesive picture of the states' positioning and final decisions emerges. The first article ("*Seeing Reason or Seeing Costs? The United States, Counterterrorism, and the Human Rights of Foreigners*") engages in theory-testing process tracing to scrutinize why the US established extraterritorial human rights safeguards during the 'War on Terror', and thus, imposed restrictions on itself. In doing so, the paper contributes to the dissertation by analyzing a powerful state's motivation behind introducing extraterritorial safeguards against torture, targeted killing, and mass surveillance even though such reforms constrained their hitherto prevailing counterterrorism strategy. The

theory-testing approach is particularly fitting as causal linkages could be deductively derived from human rights literature (see Bennet and Checkel, 2015; Van Evra, 1997: 65). Hence, the gathered data has been inspected for mechanistic evidence that supports or weakens the conjecture of three predefined causal mechanisms of social influence (*normative persuasion, strategic learning, coercion*). To increase the confidence in the resulting findings, independent mechanistic fingerprints have been aggregated throughout multiple rounds of research (Waldner, 2012). Despite the overall deductive approach, the study remained open for inductive findings regarding supporting conditions (see Beach and Pedersen, 2019: 269).

Subsequently, the second paper (*"When Identity Meets Strategy – The development of British and German anti-torture policies since 9/11"*) provides a content analysis of British and German data to analyze why the two states have taken opposite strategic decisions regarding the introduction of safeguards against torture, even if they found themselves targeted by similar pressures and measures of influence. With this research question, the article contributes to the dissertation by capturing the nuances and dynamics of strategic decision-making by showing how strategic preferences emerge and then translate into policy change or continuance. By using rationalist theory and by carving out corresponding constraints by relying on supplementary normative concepts, the analysis traces three categories of strategic preferences (*national, international, and elite*) through the interviews and further primary sources to filter out the state leaders' rationale behind their decision-making.

Finally, the third paper ("*British counterterrorism, the international prohibition of torture, and the Multiple Streams Framework"*) uses theory-building process-tracing to scrutinize on the micro-level why politicians pursue policy-change even if their re-election is not at risk and the public remains indifferent if not opposing to such changes. By zooming in on the details of one specific British policy reform, the paper contributes to the dissertation by looking into how strategic foreign policy interests translate into domestic policy-making processes. By combining the findings of the first two papers with theory from the field of policy analysis, a field which predominantly focuses on domestic influences (see e.g. Herweg et al., 2015), paper 3 also enhances the theoretical scope of the dissertation. In light of the apparent discrepancies between existing theory and empirical observations, theory-building process-tracing has been used to expand theoretical postulations of the Multiple Streams Framework to include foreign policy preferences (Geddes, 2003).

5. Research Results and Implications

The following section delves into the dissertation's three articles, as well as into their implications for the broader academic debate. In order to do so, each paper is presented individually along with the corresponding theoretical postulations and the final results. Afterwards, the papers' outcomes are combined to discuss their empirical and theoretical implications, before eventually the last sub-sections elaborate on avenues for future research. All three papers cover different theoretical facets and empirical cases, but ultimately share common notions regarding Rationalism and the violation of the right to be free from torture as an exemplification of an extraterritorial human right. In addition, however, article 1 offers complementary insights into the evolution of US safeguards against mass surveillance and targeted killings, enabling thus, tentative inferences on how protection clauses in other policy fields have emerged.

5.1. "Seeing Reason or Seeing Costs? The United States, Counterterrorism, and the Human Rights of Foreigners"
(M. Heupel, C. Heaphy, & J. Heaphy, published in: *The European Journal of International Relations*)

The first paper was inspired by an empirical contradiction, which in a similar fashion runs like a red thread throughout the dissertation: The US' decision to self-impose restrictions on their counterterrorist-strategies in order to protect foreigners abroad against American human rights violations. Since 9/11 and in the spirit of preventing any future attacks, the US military and intelligence agencies repeatedly breached the basic principles of the Convention against Torture, while also infringing on foreigners' right to privacy and engaging in an ever-expanding drone warfare. While these violations have been thoroughly scrutinized, another related phenomenon has been largely overlooked, namely the gradual introduction of corresponding extraterritorial safeguards. Hence it remains puzzling that in the midst of the 'War on Terror' and after years of violating their rights, the US would eventually restrict their own security strategy to, at least officially, grant foreigners abroad human rights protections via for instance, Executive Order 13491 which introduced a ban on torture and on cruel, inhuman, or degrading treatment. Consequently, the first paper deals in particular with the question of why the US established these extraterritorial safeguards linking as a result the research to the broader accountability debate.

In deriving three central mechanisms of social influence from existing literature, the paper explores the interrelation of micro-macro processes by tracing not only the impact of strategic learning processes among American decision-making, but also by considering normative and coercive means of social influence. While all three causal mechanisms share the common cause 'extraterritorial human rights violation', and the outcome 'safeguard' on the macro level, the intermediary steps being triggered on the micro-level differ in each case and are mutually exclusive (see figure 3).

Starting Point →	Intervention →	Processing →	Outcome
Moral persuasion			
Extraterritorial human rights violations →	Moral arguments →	Acceptance of norm →	Safeguards
Strategic learning			
Extraterritorial human rights violations →	Strategic arguments →	Anticipation of future strategic gains →	Safeguards
Coercion			
Extraterritorial human rights violations →	Material sanctions, shaming, or litigation →	Perception of an urgent need to react →	Safeguards

Figure 3: Mechanisms of Social Influence; taken from Artikel 1, p. 4.

To begin with, the mechanism of moral persuasion postulates that in light of a human rights violation, a norm entrepreneur confronts the incumbent decision-makers with their moral wrongdoings and attempts to persuade them of the norm's intrinsic value (Finnemore and Sikkink, 1998). If successful, the intervention triggers a state of cognitive dissonance within the decision-makers (see Reinold and Zürn, 2014), leading ultimately to the introduction of a new safeguard, which aligns state behavior and moral expectations. Strategic learning, in turn, also foresees an intervention, but rather on strategic grounds, meaning that before any tangible pressure builds up, internal or external actors confront the rational policymakers with strategic arguments on why their current conduct undermines their long-term goals (see Grobe, 2010). Based on these anticipatory notions of counterproductivity, incumbent politicians process the

arguments and run a cost-benefit analysis, introducing eventually a new safeguard if they perceive policy-change as more advantageous than policy-continuance. Finally, the mechanism of coercion comprises of three different types of intervention, which, however, all put immediate pressure on the incumbent decision-makers, and thus create a perceived urgent need to act, which paves the way for a new safeguard. In the case of litigation, for instance, court judgements pressure the government to alter their conduct (Duffy, 2018); shaming immediately targets the state's reputation by NGOs and the media making graphic evidence of the human rights violation accessible to the public (Krain, 2012; Murdie and Davis, 2012), while material sanctions are used by IOs and other states to exploit the perpetrating state's dependencies and punish it for its human rights violating behavior (Donno and Neureiter, 2018; Hafner-Burton, 2005).

After a thorough theory-testing process-tracing analysis, paper 1 highlights the importance of strategic and coercive measures, while finding little evidence for a successful unfolding of the normative mechanism. In addition, the study offers inductive findings regarding the causal mechanisms' scope conditions. Accordingly, in two of the three examined cases, coercion was found as the sole mechanism (right to privacy/ material sanctioning) or as one of two mechanisms (right to be free from torture/ shaming) leading to policy-change, proving thus that under certain conditions even powerful states can be held accountable by coercive measures. Although powerful states might be less likely to face rejection when attempting to join an IO or to be excluded from preferential trade agreements (see Hafner-Burton, 2008; Schimmelfennig and Sedelmeier, 2004), they can still be vulnerable to coercion if the state grows disproportionately dependent on other actors, as was the case of mass surveillance, in the US government, which relied heavily on its access to foreign citizens' data via US tech companies. Consequently, targeted sanctions against the cooperation partners can exploit this vulnerability and make policy-continuance prohibitively costly for the government. Similarly with shaming, powerful states might be able to counterweight reputational losses with material power, or the public might in general be less inclined to intervene for the rights of foreigners abroad; but shaming has still proven to be a useful mechanism if the campaign can explicitly visualize the breach of an internationally accepted taboo such as torture (Barnes, 2017; Keck and Sikkink, 1998: 205).

Similar to coercion, strategic learning has likewise impacted two incidents of policy-change; once as the sole mechanisms (right to life) and once as one out of two mechanisms (right to be free from torture). Important here is the anticipatory notion of the strategic arguments, meaning that at the point of intervention no coercive pressure was being targeted against the decision-

makers, but that rather the politicians processed the rational arguments and induced policy-change to prevent future negative costs. In both cases the credibility of the intervening actor and the actual likelihood of the negative consequences materializing constituted key factors in convincing the policymakers to consider change (see Bapat et al., 2013: 89–90; Haas, 2004). Especially regarding the first condition, credibility, it seemed crucial in both cases that the intervention was not only pushed forward by liberal NGOs, but also substantiated by other experts such as high-ranked CIA or military officials who are respected from both progressive and conservative politicians.

Finally, moral persuasion seems to have played only a minor role in the introduction of extraterritorial human rights safeguards. While this might be surprising in the context of a country that frequently identifies itself as strong human rights defender, the finding still offers interesting insights on the topic, as it substantiates claims that in general the US has only internalized human rights norms for their own citizens and that in cases of norm-conflict security related issues are given a higher priority than human rights (Sikkink, 2013). This is not to say, however, that individual actors within the government did not believe in the norm, in fact, intervening actors in the other mechanisms might very well have been motivated by their deep normative beliefs. Rather, the article claims that normative arguments have not been enough to trigger policy-change on their own. This observation can also explain why most of the safeguards do not constitute clear bans of the respective violations, as would be expected if the policymakers truly believed in the norms, but rather continue providing great levels of discrepancy and loopholes. With that being said, however, the safeguards are nonetheless important, as they show how policymakers gradually recognize the consequences of violating the rights of foreigners abroad.

5.2. "When Identity Meets Strategy – The development of British and German anti-torture policies since 9/11"
(J. Heaphy, under review in: *International Studies Perspectives*)

Inspired by the first paper, the second article delves deeper into the constraints of strategic policy-making by exploring why the UK and Germany chose opposing policy strategies when being confronted with similar amounts of pressure regarding their abetting role in the CIA Detention and Interrogation Program. During the 'War on Terror', both states assisted the US by enabling rendition flights via their respective airspaces and by using their intelligence agencies to channel vital information and question sets to their American counterparts

(European Parliament, 2006)[2]. Eventually, however, despite differing degrees of involvement, both countries faced relatively equal levels of backlash. The EU Parliament called out both states side by side for their active involvement in the CIA's detainee mistreatment (European Parliament, 2006); NGOs in both countries shamed the SIS and BND for their norm-violating behaviors (Amnesty International, 2009; Human Rights Watch, 2006), domestic parliamentary inquiries confirmed indirect violations of the CAT, and in both countries, despite individual waves of public uproar, the public pressure remained moderate. Yet, in the end, one major difference could be observed: While David Cameron eventually introduced the so-called 'Consolidated Guidance', a reform to the intelligence agencies' interrogation practices abroad, Germany opted rather for policy-continuance.

Even though Cameron's policy does not constitute a panacea against torture and in fact it entails major loopholes, it still poses an interesting puzzle: Despite many contextual similarities and shared Western democratic values, the decision-makers ultimately came to different conclusions on whether it was strategically beneficial for their states to pursue policy-change. Hence, the second article explores why the two states embarked on different policy trajectories, even if they found themselves in contextually similar situations and encountered akin means of pressure and influence. In order to enable a detailed analysis, the article only focuses on policies regarding the states' intelligence agencies and is limited to the time between 2001 and 2010, excluding thus compounding factors such as Brexit and leaks regarding previous British-Libyan intelligence cooperation.

To capture the states' reasoning behind their decision-making, the article follows rationalist theory, while considering different notions of state identity to fully capture the nuances of and constraints to the states' strategic preferences. Hence, the article assumes British and German policymakers to be rational actors, who base their decisions on cost-benefit analyses in order to maximize their interactions' gains and to avoid costly consequences (Snidal, 2002). A purely rationalist analysis of the empirics reflects such notions indeed, but provides little insight into why the two states' strategic calculations differ so greatly. After all, Rationalism does certainly account for heterogeneous risk, action, and outcome preferences (Fearon and Wendt 2002, 59), but still falls short in fully explaining why and how these differences emerge.

Therefore, the articles takes inspiration from different facets of state identity – namely national, international, and political elite identity - to deduct country-specific constraints limiting the

---

[2] In the second article, the focus is laid on human rights violations and policy-changes regarding the British and German intelligence services, not the military.

respective decision-makers options and strategic choices. To begin, the concept of *national identity* alludes to a country's history, traditions, norms and values, as well as the country's distinctive political culture (Gillis, 1994; Weedon and Jordan, 2012). The resulting shared sense of cohesiveness and collective national memory among the respective state's citizens ultimately marks the limits and boundaries of what the people perceive as (un-)desirable political actions at home and abroad (Ryan, 2012). The state's *international identity*, however, encompasses a country's self-image on the global stage, its reputation among other international actors, respective dependencies, and power-relations as well as the state's legitimacy claims for its international actions (Manners and Whitman, 2003: 383). Hence, respective preferences and constraints emerge from international regimes, alliances, rivalries, and other power structures, which dictate the consequences of the states' actions. Finally, the *political elite identity* accounts for government constellations, dominating party ideologies, and the notion of culpability within the governing body, deriving thus potential identity-based constraints from which politicians and parties are in power, what the incumbent government's main policy goals are, and who was to take the blame for current or previous wrongdoings (Weaver, 1986: 371; Raunio and Wagner, 2020: 515).

When applying these theoretical facets of identity to the empirical cases, different constraints in the British and German decision-making become apparent, as for instance aspects of the British national identity supported policy-change, while German national identity deemed potential alterations costly. Accordingly, the clear parallels between the British assistance to the CIA and the UK's use of the 'Five Techniques' in Northern Ireland[3] made it imperative for the British government to convey a message of taking any torture-related allegations seriously in order to prevent further distrust against the government spreading in Northern Ireland, and to thus preempt a destabilization of domestic relations. Furthermore, British policymakers worried that change was needed to domestically protect the reputation and the morale of the British intelligence community. In contrast to the UK, German history had a considerably different impact, as it had to be expected that any policy-change could be construed as an admission of guilt, which in the context of torture and against the background of WWII would carry very high political costs that risked inciting public indignity. Similarly, the absence of public concern on the issue failed to justify the risk of being accused of failing to regulate the

---

[3] 'The 5 Techniques' refer to five torture techniques (prolonged wall-standing, hooding, subjection to noise, deprivation of sleep, and deprivation of food and drink), which the British used against the Irish in the Northern Ireland Conflict and officially prohibited in the aftermath. All five practices have, however, been later reflected in the American Enhanced Interrogation Techniques.

country's security agencies, a notion which given Germany's de-militarized political culture could likely be particularly costly.

In the context of the states' international identities, parallel patterns of constraint can be observed; in fact, the UK pursued change to uphold the country's international image, while Germany opted for policy-continuance for the very same reasons. The British decision was not only important on the British domestic stage, but also internationally in order to bolster and restore the UK's carefully crafted self-image of a 'beacon of the rule of law' and democratic leader in the world. In line with this, the allegations likewise jeopardized the UK's reputation, meaning the image other countries had of the UK, and thus future intelligence cooperation as well as the future of the 145 British military sites abroad. Contrastingly, a key factor opposing change in Germany was its high dependency on the US in terms of security. In the absence of similarly recent military scandals as was the case for the UK, German decision-makers feared that pursuing policy change would put the spotlight on human rights violations resulting from American-German cooperation, undermining thus the tediously crafted image of German loyalty towards the US. Additionally, German policymakers feared that a new policy would introduce new hurdles for future intelligence cooperation, and thus not only tarnish its international reputation but also generate substantive strategic long-time costs. Having already earned the reputation of band-wagoning on the efforts of other NATO members, policies further complicating international cooperation, could aggravate Germany's bureaucratic and passive reputation among its allies, making the country a less attractive security partner.

Finally, the pattern of German constraints hindering policy-change and British constraints pushing change also runs like a red thread throughout the *political elite identity*. In terms of British domestic political power constellations, the desire to differentiate Cameron's newly elected administration from the previous Labour governments promised strategic benefits; especially as neither retribution nor increased notions of culpability were to be expected. In addition, a policy reform also promised to aid relations within the young coalition with the Liberal Democrats, who were the only party that included investigations and policy-change regarding the torture allegations into their party manifesto. In contrast, the German government constellation had not changed since 2005, meaning that Angela Merkel had been chancellor for a large share of time during which the indirect CAT violations took place, creating thus a high risk of political costs due to direct culpability and blame attribution. Furthermore, when looking at the governing parties' election campaigns, the CDU/CSU and the SPD had clearly emphasized national security concerns over the prohibition of torture, confirming repeatedly that torture, as last resort, would be legitimate in a ticking-time-bomb-scenario.

In sum, the second paper shows how multi-level constraints have shaped the countries' strategic preferences and thus determined the states' differing reactions to the gradually increasing pressure. Hence, the empirical study shows, how equal means of pressure and social influence can lead to different outcomes, emphasizing thus the importance of well-informed and country-specific argumentative interventions. In the same context, the article concludes that the interventions in the British case proved to be successful as they threatened to undermine the UK's (inter)national image and legitimacy, resulting ultimately in policy-change due to long-term cost evasion, rather than due to intrinsic normative beliefs. While this strategic commitment ultimately translated into a window-dressing policy introduction (see Hafner-Burton and Tsuitsui, 2005: 1378), it nonetheless constituted an important step, as it shows that states like the UK can be pressured into action, where small concessions can eventually lead to actual change (see Risse et al., 1998); a step that Germany, who has not changed its respective policies since 2001, has not taken yet.

5.3. "British counterterrorism, the international prohibition of torture, and the Multiple Streams Framework"

(J. Heaphy, published in: *Policy & Politics*)

Whereas the first article shows the relevance of strategic learning and the second article zooms in on the composition of states' strategic preferences, the third article finally traces how the decision-makers' choice to follow strategic interests pertaining to the international level translates to a domestic stage that is neutral, if not opposing to such policy-changes. It is especially this public indifference, which poses an interesting theoretical puzzle, given that the Multiple Streams Framework, a leading theoretical model in the field of policy analysis, claims that *"[t]he more a condition puts the policy makers' re-election at risk, the more likely it is to open a policy window in the problem stream"* (Herweg et al., 2015: 437). Hence, the third article delves into the policy-making process behind the second British reform of interrogation procedures for foreigners abroad, the so-called "Principles"[4], using theory-building process-tracing to answer the following question: How can we explain the opening of problem windows, when the triggering conditions do *not* put the decision-makers' re-election at risk? The UK provides in this context a particularly interesting case; by pursuing policy-change, May risked upsetting her voters significantly more than if she had opted for policy-continuance.

---

[4] "Principles relating to the detention and interviewing of detainees overseas and the passing and receipt of intelligence relating to detainees" (HM Government, 2019; henceforth "Principles").

Accordingly, the "Principles" were not only passed amidst Brexit negotiations, but also in face of open resistance from the public to any further investigations into misconduct by the military and intelligence agencies, during which human rights lawyers were frequently condemned in political debates as left-wing activists, who "harangue and harass the bravest of the brave [the British military]" (May, 2016).

As alluded to in the research puzzle, the third article builds on the basic concepts of the Multiple Streams Framework (MSF), a theoretical model typically used to factor domestic dynamics into policy-making processes, but which expanded to also account for international stimuli for change throughout the article. In general, MSF maintains that policy-making procedures are not inherently rational and linear processes, but rather the product of contingent stream development and a policy entrepreneur's successful exploitation of a policy window (Herweg et al., 2015). In detail, this means that the problem, political, and policy stream develop independently from each other until a policy window, a fleeting opportunity for advocates to push for change, opens and enables the alignment of the three streams, resulting thus, in a policy-proposal being adopted to the decision-agenda. In this context, the problem stream encompasses attention-seeking behavior by policy entrepreneurs, who exploit focusing events, indicators, or negative feedback in order to create a sense of imminent pressure, which shall compel the incumbent decision-makers to take action (Brunner, 2008: 52). The political stream, in contrast, accounts for the impact that the general public mood, electoral turnovers, or a politician's ideology has on prevailing political dynamics and the decision-makers' disposition for change (Cairney and Zahariadis, 2016: 93). The policy stream, in turn, rather focuses on the development of feasible policy alternatives delineating how ideas emerge within the respective policy communities, where they are evaluated, modified, and combined until they are narrowed down to a short-list of feasible policy options (Zhu, 2008: 317). In order for policy-change to occur, a policy entrepreneur must invest their resources and skills to successfully align all three streams so that a fully developed policy alternative is perceived as matching the politicians' goals and as countering an at that time pressing problem (Kingdon, 1984: 19).

In light of this, MSF foresees three causal mechanisms (CM) of how a policy entrepreneur can align the different streams (see Fig. 4). The first mechanism, *doctrinal coupling*, accounts for processes in which an elected official, uses policy windows like electoral turnovers or public mood swings to push for a new policy alternative by evoking a hitherto unrelated, but fitting narrative from the problem stream to gain broad support for their policy claim (Boscarino, 2009: 416; Herweg et al., 2018: 27). *Consequential coupling*, in contrast, describes how policy entrepreneurs frame policy windows like a focusing event, or negative feedback to convince

decision-makers of adapting their problem agenda (Zahariadis, 2003: 72) and to subsequently sell their policy alternatives as best option available (Jones, 2003: 396). Lastly, *spillover commissioning*, outlines how institutional spillovers from one policy field to another open a policy window in the politics stream by putting the affected officials under such a reform pressure that they actively seek the help of the policy community (Ackrill and Kay, 2011).

Cause		Policy Window		Stream Alignment		Outcome
Doctrinal Coupling						
Contingent Stream Development	→	Window opens in the Political Stream enabling politician(s) to promote their pet policy	→	Politician(s) engages in problem-oriented alliance-building	→	Policy adoption
Consequential Coupling						
Contingent Stream Development	→	Window opens in the Problem Stream compelling politician(s) to find a quick solution	→	Policy entrepreneur(s) targets politician(s) for policy-selling	→	Policy adoption
Spillover Commissioning						
Contingent Stream Development	→	Window opens in the Political Stream forcing politician(s) to manage institutional spillovers	→	Politician(s) selects already known policy proposal	→	Policy adoption

Figure 4: Mechanisms of Stream Alignment within the Multiple Streams Framework; taken from Artikel 3, p.6.

Regarding a central pre-condition for consequential coupling, current MSF literature states that *"[t]he more a condition puts the policy makers' re-election at risk, the more likely it is to open a policy window in the problem stream"* (Herweg et al., 2015: 437), this hypothesis does not, however, match with the empirical findings in the British case. Effectively Herweg et al.'s hypothesis means, that politicians will pursue policy-change if a crisis, negative economic indicators, or critical feedback by parliamentary investigations, for instance, threaten their popularity with the voters. While these postulations have repeatedly been applied to the domestic level (see e.g. Zohlnhöfer, 2016; Cairney, 2018), there still exists a mismatch between theory and empirics when looking at the policy-making process behind the British "Principles". Accordingly, procedures of *spillover commissioning* can be ruled out in the British case as there

have not been any institutional spillovers. Similarly, *doctrinal coupling* is not applicable as at that time, neither the public mood nor the Conservative party's ideology highly prioritized human rights protections for foreigners abroad (May, 2016), let alone did they have a respective pet policy they wanted to promote. Finally, *consequential coupling* does seem to fit at least broadly, as after all, there has been negative feedback given to the government by the parliament (ISC, 2018a&b) and policy entrepreneurs did seek to sell their policy alternatives via the Investigatory Powers Commissioner's Office regulated procedures. However, the parliamentary investigation failed to put electoral pressure on Theresa May, whose strong disposition to pursue change anyways showcases underlying motivations, which hitherto could not be captured by the MSF model.

In order to address this incongruity, findings abstracted from an in-depth analysis of the Principle's evolution have been used to refine and expand Herweg et al's hypothesis. Accordingly, the problem window did not open in the British case because of electoral pressures, but rather because the ISC reports clearly stated that without policy-change the UK's reputation and credibility would be damaged in the long-term and thwart the state's influence on global governance structures (ISC, 2018a&b). Hence, if following the essence of Herweg et al.'s (2015: 437) postulation, namely that politicians are inherently interested in maintaining their political power and influence, the international notions derived from the empirical case do not contradict the intrinsic logic of the original hypothesis. In the end, by losing the country's access to the international negotiation table, the decrease in state power could directly translate into a decrease of the leading politicians' individual power. Whether it is having a direct influence on how international law is determined or having the capacity to promote and enforce it abroad, politicians both as individuals and representatives of the state have a compelling interest in maintaining their country's international power – without access to the negotiation table, the politician's voice will likewise remain unheard.

Building on these inferences, the article proposes a refined hypothesis for the opening of problem windows, which can capture how domestic as well as international issue perceptions can affect the politicians' basic preference of power retention and thus motivate them to adapt their problem agenda accordingly:

H: The more a condition puts the policymakers' influence on negotiations and general decision-making processes at risk, the more likely it is to open a policy window in the problem stream.

In sum, this hypothesis argues that an issue passes the threshold of opening a problem window when one or more politicians fear to forfeit their say in negotiating generally binding decisions

– regardless whether these decisions are being made on the domestic or the international stage. Hence, if applied to the domestic realm, this risk ensues if the politicians' re-election and thus their political position is endangered (see Herweg et al., 2015); in the international realm, however, the risk of losing their say emerges if their personal reputation or the reputation of the state as whole is damaged. Such a loss of credibility could restrict the politicians' access to international negotiation rounds or, for instance, impede their ability to hold other states accountable.

In addition to these theoretical findings, however, the empirical analysis also allows for further inferences on Brexit's facilitating role in the policy-change and May's pro-active policy commissioning. Accordingly, the UK leaving the European Union bolstered May's pursuit of policy-change in two significant ways: Firstly, it dominated most political debates at that time, distracting thus the public from parallel policy processes, and providing the decision-makers with sufficient leeway to pursue contentious policy reforms. Secondly, it increased the UK's dependence on its international reputation even more, as new partnerships had to be negotiated, while simultaneously trying to limit the diplomatic damages with the European neighbors. In a similar fashion, May avoided potential electoral pushbacks by diminishing her role in the policy drafting process and by shifting the drafting responsibility towards the Investigatory Powers Commissioner's Office, a political move which, given the commissions expertise, also increased the new policy's legitimacy. More importantly, however, by assigning Investigatory Powers Commissioner's Office, May could pre-empt the appointment of a less government-friendly commission or policy entrepreneur, which ultimately enabled her to address the international predicaments, while safeguarding the consideration of the government's operational interests in the new policy draft (IPCO, 2021).

5.4. *General Theoretical and Empirical Implications*

When combining the results of the individual articles, there are four prominent findings of this cumulative thesis, each of which answers one of the guiding questions and themes, while contributing to various discussions within International Relations research on human rights, norm compliance and commitment, as well as on corresponding debates centering on accountability. In the following section, the findings and implications gained from the CAT-related policy developments will be thoroughly discussed in the general context of extraterritorial human rights violations and accordingly linked to the broader academic debate as well as future avenues for research.

Finding I: Strategic calculations motivate policymakers

Although the three articles examine different facets of policy changes across different countries in the global war on terrorism, the motivations behind each of the examined policy changes shared a notable characteristic: Namely, strategic calculations played a significant role in considerations by the US, UK, and Germany, on whether to recognize and or advance human rights obligations to foreigners abroad. Within these calculations, states assess and act according to their own interests, thus following individual preferences when considering the advantages and disadvantages of potential policies. The first finding of this thesis therefore reflects rationalist theoretical patterns within the existing literature (see Epstein, 2013: 293; Fearon, 1995; Goldsmith and Posner, 2005). Additionally, however, the articles take the research further by not only considering what those interests are in the context of the post-9/11 security frameworks of the US, UK, and Germany, but also by outlining how actors from inside or outside the government can manipulate these interests, if they bear in mind country-specific parameters like state identity and particular power constellations.

In the case of the US, for example, norm entrepreneurs alongside prominent actors within the government were able to coordinate efforts to advance alternative perspectives within interrogation policy debates. Utilizing their shared expertise as human rights advocates and military strategists, a campaign headed by Human Rights First was able to align key US objectives, namely the collection of credible intelligence during military and CIA interrogations with the advancement of safeguards for foreigners captured abroad. By ultimately capitalizing on the frustration of US interrogators who were concerned that unnecessary intelligence was being lost through the lack of trust created by excessive use of fear and force, the campaign was able to gain influential allies who were respected for their seniority in interrogation operations and their political power in the American system. Together, they argued that not only would such safeguards *not* hinder efforts to obtain critical information from potential terrorists, but that it would likely advance such intelligence operations by gaining trust and undermining terrorist recruitment efforts. As a result, they successfully convinced sufficient actors within the US government that torture was in fact counterintuitive to long-term counter-terrorism efforts, thereby altering the interests and calculations of US policymakers.

As shown above and within the dissertation's other articles, such an invocation of anticipatory costs is of particular relevance when wanting to trigger strategic learning processes among decision-makers. Accordingly, it is not only the intervening actor's credibility that is crucial in facilitating learning, but also the intervention's timeframe. Hence, imminent strategic costs, even if not brought to light by shaming campaigns, can trigger politicians to behave defensively

leading thus to premature, imprudent, or evasive policies that either focus on distraction rather than on problem solving, or do not thoroughly consider all consequences of the respective policy-change. Especially the latter is for instance a frequent source of criticism regarding the current American ban on torture as it binds the CIA to an army field manual's interrogation guidelines that were originally written for low-ranked soldiers on the ground, rather than for senior interrogators specialized on intelligence collection. Hence, if in turn the intervening actors does, however, present anticipated costs, the pressure for policy-change is still high, but leaves the politicians the time to actually process the presented arguments, consider policy options, and eventually learn from the intervention.

Finding II: The effectiveness of strategic learning and shaming in the context of powerful states

The second prominent finding of this thesis is that although both shaming and strategic-learning were important mechanisms to facilitate policy-changes and to hold the perpetrating states accountable, as claimed by current literature (see Cardenas, 2004), corresponding concepts and mechanisms have to be refined in the context of powerful states. Accordingly, most existing literature on norm-compliance and commitment focuses predominantly on transitioning states, with only very few exceptions looking at backsliding behavior by previously "socialized" (see Risse et al., 1998) Western democracies. Hence, the mechanism of strategic learning might overall be applicable, but important differences in the conceptualization and specification have to be considered. For instance, existing literature suggests manipulating a state's cost-benefit-analysis by incentivizing commitment with access to certain regional or international organizations, with preferential trade agreements, or with foreign investments (Elliot, 2018). In the cases of the US, the UK, and Germany, however, the effect of such incentives is largely irrelevant given their place in the international order since all three states already have access to these benefits. Instead, impactful argumentation in those cases has predominantly centered on efficiency and the maintenance of legitimacy. While the first, can certainly be directly translated into the setting of transitional states, the latter is rather unique to powerful states, who are not just starting to build a norm-compliant reputation for themselves, but who, especially in the case of Western democracies, justify their actions and prominent position on the international stage with the carefully crafted image of being human rights respecting and international law-abiding state actors.

This incentive of maintaining legitimacy goes hand in hand with another precondition for strategic learning among powerful states: the ability to pursue change without losing face.

Accordingly, for most transitional states the official commitment to human rights regimes often translates directly into an increase in legitimacy (Vreeland, 2008). In the case of countries who display backsliding behavior however, a respective policy-change is often perceived as an open admission of guilt damaging thus the state's reputation first before then potentially restoring it. Hence, if strategic arguments carefully embed this differing nuance in their rationale and as a result provide an option for policy-change that does not force the state actors to admit culpability, as was a major concern in the German case, the likelihood of success increases. Particularly important in this endeavor are anticipatory notions of looming pressures, meaning that intervening actors stress the salience of a policy's long-term negative consequences, while, nonetheless, granting the targeted decision-makers the possibility to look for fitting policy-alternatives without immediately triggering defensive behavioral patterns. While such an approach might bare the risks of window-dressing policy-making the ensuing policies might still have their own value in constituting small, incremental steps towards full commitment and opening the field to rhetorical entrapments (see Schimmelfennig, 2001); a further discussion on window-dressing will be provided in the section on finding IV.

When looking at the mechanism of shaming, a particular characteristic of powerful states likewise impacts the process' dynamics, namely their ability to avoid accountability. Shaming, when successful, can jeopardize a states' carefully crafted image on the international stage in the long-term, thereby amplifying pressure for change. Directly shaming a state for its actions or the consequences of which, however, can be particularly difficult due to states' various evasion strategies (see e.g. Adler-Nissen, 2014; Mitchell, 2012; Mor, 2009). Accordingly, targeted states can invoke narratives of denial, as seen in the British case (Blakeley and Raphael, 2020), or deviate a problem, as seen in Germany's reaction to torture allegations. In the end, irrespective of the exact method of evasion, the current academic debate provides a plethora of possibilities of how states can attempt to counter shaming campaigns. While such attempts are not unique to powerful states, the reach and impact of such evasion campaigns grows in parallel with a state's rank in the international order since power is among others defined as the ability to shape narratives and to force another actor to do something they otherwise would not do (Dahl, 1957: 202). Hence, while transitioning states might invoke evading narratives, powerful states can furthermore pressure other actors, under threat of ending preferential treatments or development aid, for instance, into echoing and reinforcing such narratives (see e.g. Baldwin, 2016: 170) enlarging thus, not only the audience, but via the increasing range of multiplicators, also the supposed credibility of such claims.

Finding III: Strategic learning and shaming in the context of direct and indirect human rights violations

While the degree of a human rights violation, meaning in the context of this study whether the states directly (US) or predominantly indirectly (UK and Germany) breached international law, does not seem to greatly affect the principles of strategic learning, it does indeed impact the functioning of shaming campaigns. When it comes to strategic learning, the main goal is to convince the decision-makers of their behavior's negative consequences by credibly outlining the anticipated costs of policy-continuance. Whereas in case of direct violations stronger statements regarding legal or reputational consequences can be made, the success of such campaigns ultimately depends on the intervening actor's skills of adapting their argumentation to the perpetrating state's situation and interests (see e.g. article 2). Hence, the preparation of an intervention might need more time and elaboration in case of indirect violations, however, the actual process of manipulating the targeted actor's cost-benefit analysis does not change.

In contrast, existing literature on shaming suggests that respective interventions are only feasible when there is clear, graphic, and unambiguous evidence that can credibly expose a state's norm-violating behavior (see Sikkink, 2013). In the case of indirect human rights violations, however, this is often not feasible. Few shaming campaigns were as successful as the case of the US and Abu Ghraib. Due to the complexity in assigning culpability to the state however, similar shaming campaigns in the UK and Germany were unable to create such a significant policy impact. Although German and British support was key to the "success" of the CIA Detention and Interrogation program, it was more difficult to link tangible evidence of the UK and Germany's less direct involvement. As a result, both countries, but Germany in particular, were able to invoke more effective deviation narratives by pointing to the relative insignificance of their "indirect" violations in comparison to the severity of US violations. Hence, it becomes apparent that for shaming indirect perpetrators new strategies are needed, while the elaboration of such lays beyond the scope of the thesis, article 2 has shown the significant impact the jeopardy of a state's international identity has. Thus, shifting the focus from graphic evidence of the violation, towards directly targeting the pillars of a particular state's identity could enable progress in respective campaigns. In this regard, articles 2 and 3 have also shown that for policy-change to occur public pressure is not necessarily imperative. Hence, instead of attempting to turn the German public against the government, for instance, it might be worth exploring the state's vulnerabilities and dependencies towards the US, an actor who already knows first-hand what had happened, adopting consequently a campaign which

rather than deteriorating the public's trust, aims at decreasing the US' officials trust in their German counterparts.

Finding IV: Norms still matter

Despite the clear challenges facing states' commitment to human rights, this thesis has demonstrated the persistent relevance of norms in three different ways. To begin with, the existence of norms has remained relevant for the initiation and facilitation of policy debates across each of the case studies even though normative persuasion (see Finnemore and Sikkink, 1998) and immediate shaming (see Krain, 2012; Murdie and Davis, 2012) have had limited success, with the exception of a few cases (e.g. Abu Ghraib). Specifically, norms motivate individual politicians and norm entrepreneurs to engage in lobbying and policy-changes. Secondly, despite such normative beliefs and corresponding persuasion campaigns having a limited impact on backsliding signatory parties to human rights, normative notions are still indispensable for increasing strategic arguments' potency by tailoring respective lines of reasoning to the targeted country's state identity (see article 2). Finally, in most strategic analyses, "costs" were frequently understood also as normative costs, and not simply as transaction costs. Hence decision-makers in the US, the UK, and Germany grew more worried about potential reputational damages or how openly norm-violating behavior might negatively affect other countries' willingness to cooperate, rather than pondering on the financial or administrative tolls of a policy-change.

Although the American and British reforms were hardly a panacea against the use of torture in their military and interrogation policies, they were nonetheless improvements to the status quo that had flourished under the Bush and Blair administrations. Despite their differing degrees of increased commitment, both states' policy-changes shared concerning loopholes and shortcomings that will likely remain obstacles to future commitment measures. These limits, however, are perhaps unsurprising given that the strategic motivations for the policy-changes were primarily to appear "compliant" (e.g. Sikkink, 2013) without necessarily committing to international law for its own sake. "Compliance" in this sense, was primarily a means to reach an advantageous end; an important limitation that continues to frustrate human rights debates (see Hafner-Burton and Tsuitsui, 2005; Koh, 2018; Lin, 2010).

The US and UK's imperfect commitment to international law nonetheless produced significant policy-changes that resulted in increased levels of transparency. These small steps remain crucial wins in comparison to the policy stagnation that occurred in the German case. As states

recommit themselves to certain norms and safeguards, they increase the political costs of future norm-violations. Additionally "updating" norm relevance in modern contexts undermines future efforts to argue that state commitment is disconnected from contemporary security dilemmas. The acknowledgement and advancement of the strategic incentives of commitment therefore grants future campaigns a relative advantage over previous efforts to hold states accountable. Furthermore, even small reforms can provide a feedback loop for compliance research and accountability efforts. Eventually, policymakers are forced to ask themselves at what point do successive evasion efforts become more costly than full commitment. This thesis thus illustrates how "window-dressing" (see Hafner-Burton and Tsuitsui, 2005) can have incremental progress in the long arc of human rights accountability.

5.5. *Avenues for Future Research*

In conclusion of this thesis, the following section outlines four avenues for future research: A mixed method study comprising a large n-sample to shine further light on current trends in accountability, a qualitative study of new developments in the anti-torture debate after 2019, further investigations into policy and accountabituy debates pertaining to joint huamn rights violations by International Ogranizations and respective member states, and finally a similar study on the potential development of extraterritorial protection mechanisms under authoritarian regimes. All suggestions are based on the thesis findings.

By zooming into CAT-related violations by the US, the UK, and Germany the thesis relies on a small-n sample, which enables an empirically rich analysis under the caveat of limited generalizability. Given the focus on three countries and a total of five policy-changes, the case selection allowed for an in-depth investigation of primary sources and the generation of new empirical findings via 43 semi-strcutured interviews with key stakeholders, producing thus not only findings which further refined existing theoretical concepts, but enhanced the academic debate with new accumulative inferences on extraterritorial policy developments in Western Democracies. In order to further augment the scholarly understanding for other policy fields and the different nuances of accountability means, additional mixed methods studies with large n-samples could shine further light on general trends in norm-compliant behavior and accountability processes. Hence, if taking Simmon's quantiative work from 2013 as a model, a quantitative study testing the success of different mechanisms of accountability and social influence across a wide array of countries, would enable scholars to better understand how certain regime types, alliances and rivalries, or international power changes influence specific

means of accountability. Building on such findings, further qualitative investigations could then shine light on the processes of shaming indirect perpetrators, a current gap in the academic debate as shown under finding II, or on the success of country-specific strategic interventions targeting in particular interests derived from the states' identity rather than employing a one-size-fits-all approach. In general, the War on Terror with torture, arbitrary detentions, renditions, and other related human rights violations would still provide a good basis for such a research as through the US' broadly cast web of secret detention facilities the behavior of roughly 55 different countries (ranging among others from EU members, to Thailand, to North Africa, and Pakistan (see Horowitz and Cammarano, 2013) could be explored while holding the independent variable of the specific human rights violation constant.

In a similar fashion, the thesis spans almost two decades, invesitgating all three countries' relevant policy-making processes between 2001 and 2019, yet, since then important changes on the states' political stages have unfolded making a follow-up study on the states' torture-debates an interesting avenue for future research. Accordingly, this dissertation found Angela Merkel's (CDU/CSU) long-lasting chancelorship to be an inhibiting factor for German policy-change as it would have equaled an admission of guilt and thus would have carried high political costs. Since the conclusion of the dissertation's analysis, however, a new government headed by Olaf Scholz (SPD) came into power changing the dynamics of German policy-making, in particular when considering the Green coalition partner, who is known to be a strong advocate for the universality of human rights. Likewise, Brexit, in the analysis a distracting parallel event supporting policy-change, has since been concluded forcing the UK to build up a new security strategy outside of the EU. In the US, in turn, President Biden, the former Vice-President under the Obama administration, which issued two out of the three analysed anti-torutre provisions, took office after four years of Donald Trump's administration, which publicly advocated for the benefits of torture. So far, the author of the dissertation is not familiar of new relevant policy-changes, however the alterations on the countries' political stages should offer grounds for interesting shifts in the respective torture debates.

Furthermore, the dissertation focuses in particular on torture-related extraterritorial human rights violations committed by states, contributing thus to a discussion that despite 9/11 being twenty years in the past, still is current and of high practical relevance. In light of an ever-increasing organizational density, future research could provide interesting insights into not only the policy dynamics behind human rights violations that were committed by member states, but those that were enabled by regional organizations. Hence, when looking, for instance, at recent developments at the EU's outer borders with member states using force, tear gas, and

sound canons to prevent migrants from entering their territory, in-depth scrutiny is needed to first disentangle the interwoven processes of EU funding and member states norm-violating behavior, to then be able to develop sound strategies of holding both actors accountable. Here, insights from the dissertations' findings could provide initial guidance as they engage with the concepts of indirect and directly violating behavior as well as with the need for specifically tailored interventions targeting a states' identity; a concept also fit to capture an organization's core values and purposes.

Finally, the findings in this dissertation identitfy a potential avenue for accountability that extends beyond Western democracies. Whereas tradtional causal mechansims, such as shaming or persuasion have proven challenging in more authoritarian regimes, which suppress investigative journalism, protests, public debates, and judicial interventions, such regimes are hardly immune to strategic reasoning. Future research that focuses on how human rights protection can be successfully alligned with the long-term interests of diverse political actors as well as understanding the factors and conditions that influence their decision-making could provide further insight into not only the accountability of "back-sliding democracies" but also that of non-democratic regimes who otherwise remain opposed to so-called Western human rights norms.

6. Bibliography

Ackrill R and Kay A (2011) Multiple streams in EU policy-making: the case of the 2005 sugar reform. *Journal of European Public Policy* 18(1): 72-89.

Adler-Nissen R (2014) Stigma management in international relations: Transgressive identities, norms, and order in international society. *International Organization*, 68(1): 143-176.

Alden C (2017) Critiques of the rational actor model and foreign policy decision making. In: Thompson WR and Capelos T (eds) *Oxford Research Encyclopedia of Politics*. Oxford: Oxford University Press.

Allendoerfer MG, Murdie A and Welch RM (2020) The Path of the Boomerang: Human Rights Campaigns, Third-Party Pressure, and Human Rights. *International Studies Quarterly* 64(1): 111-119.

Amnesty International (2009) Gegen den Strom. Available at: https://www.amnesty.de/journal/2009/juni/gegen-den-strom (accessed 15 September 2021).

Andreeva C (2021) The evolution of information-sharing in EU counter-terrorism post-2015: a paradigm shift? *Global Affairs* 7(5): 1-26.

Arsenault E (2017) *How the gloves came off: Lawyers, policy makers, and norms in the debate on torture*. New York: Columbia University Press.

Auswaertiges Amt (2022) Human Rights. Available at: https://www.auswaertiges-amt.de/en/aussenpolitik/themen/menschenrechte (accessed 06 September 2021).

Baldwin DA (2016) *Power and international relations: A conceptual approach*. Princeton: Princeton University Press.

Bapat NA, Heinrich T, Kobayashi Y and Morgan TC (2013) Determinants of sanctions effectiveness: Sensitivity analysis using new data. *International Interactions* 39(1): 79-98.

Barnes J (2017) *A Genealogy of the Torture Taboo*. London: Routledge.

Bassano D (2014) The Boomerang Pattern: Verification and Modification. *Peace & Change*, 39(1): 23-48.

Bates ES (2014) Sophisticated constructivism in human rights compliance theory. *European Journal of International Law* 25(4): 1169-1182.

Beach D and Pedersen RB (2019) *Process-tracing methods: Foundations and guidelines*. Ann Arbor: University of Michigan Press.

Bennett A and Checkel JT (2015) Process tracing: From philosophical roots to best practices. In: Bennett A and Checkel JT (eds) *Process tracing: From metaphor to analytic tool*. Cambridge: Cambridge University Press: pp. 3-37.

Bennett A and Elman C (2006) Qualitative research: Recent developments in case study methods. *Annual Review of Political Science* 9(1): 455-476.

Bennett A and George AL (1997) *Process tracing in case study research*. Washington, DC: MacArthur Program on Case Studies.

Blakeley R and Raphael S (2017) British torture in the 'war on terror'. *European Journal of International Relations* 23(2): 243-266.

Blakeley R and Raphael S (2020) Accountability, denial and the future-proofing of British torture. *International Affairs* 96(3): 691-709.

Boscarino JE (2009) Surfing for problems: Advocacy group strategy in US forestry policy. *Policy Studies Journal* 37(3): 415-434.

Brimbal L, Kleinman SM, Oleszkiewicz S and Meissner CA (2019) Developing rapport and trust in the interrogative context: An empirically-supported and ethic alternative to customary interrogation practices. In: Barela SJ, Fallon MJ, Gaggioli G, Ohlin JD (eds) *Interrogation and Torture*: *Integrating efficacy with law and morality*. Oxford: Oxford University Press, pp. 141-170.

Brunner S (2008) Understanding policy change: Multiple streams and emissions trading in Germany. *Global Environmental Change* 18(3): 501-507.

Cairney P and Zahariadis N (2016) Multiple streams approach: a flexible metaphor presents an opportunity to operationalize agenda setting processes. In: Zahariadis N and Buckman M (eds) *Handbook of Public Policy Agenda Setting*. Cheltenham: Edward Elgar Publishing: 87-104.

Campbell JE (1992) Forecasting the presidential vote in the states. *American Journal of Political Science* 36(2): 386-407.

Cardenas S (2004) Norm collision: Explaining the effects of international human rights pressure on state behavior. *International Studies Review* 6(2): 213-231.

Cardenas S (2007) Conflict and Compliance State Responses to International Human Rights Pressure. Philadelphia: University of Pennsylvania Press.

Cairney P (2018) Three habits of successful policy entrepreneurs. *Policy & Politics* 46(2): 199-215.

Clapham A (2006) Human rights obligations of non-state actors in conflict situations. *International Review of the Red Cross* 88(863): 491-523.

Clifford GM (2017) Just Counterterrorism. *Critical Studies on Terrorism* 10 (1): 67–92.

Cole WM (2009) Hard and Soft Commitments to Human Rights Treaties, 1966–2000. *Sociological Forum* 24(3): 563-588.

Coates AJ (2016) *The ethics of war*. Manchester: Manchester University Press.

Davenport C (2000) *Paths to State Repression: Human Rights Violations and Contentious Politics*. Lanham: Rowman & Littlefield.

Davies WA (2018) Counterterrorism effectiveness to jihadists in Western Europe and the United States: we are losing the war on terror. *Studies in Conflict & Terrorism* 41(4): 281-296.

Dahl RA (1957) The concept of power. *Behavioral Science* 2(3): 201-215.

Des Forger A (1999) *Leave None to Tell the Story: Genocide in Rwanda*. New York: Human Rights Watch.

Deutscher Bundestag (2009) Beschlussempfehlung und Bericht des 1. Untersuchungsausschusses nach Artikel 44 des Grundgesetzes. Available at: https://dserver.bundestag.de/btd/16/134/1613400.pdf (accessed 03 August 2021).

Donnelly J(1989) *Universal Human Rights in Theory and Practice*. Ithaca, NY: Cornell University Press.

Donno D and Neureiter M (2018) Can human rights conditionality reduce repression? Examining the European Union's economic agreements. *The Review of International Organizations* 13(3): 335-357.

Duffield John (1998) *World power forsaken: political culture, international institutions, and German security policy after unification*. Palo Alto: Stanford University Press.

Duffy H (2018) *Strategic human rights litigation: Understanding and maximising impact*. Oxford: Bloomsbury Publishing.

Elliott KA (2018) Factors affecting the success of sanctions. In: Cortright D and Lopez GA (eds) *Economic sanctions: Panacea or peacebuilding in a post-cold war world?* New York: Routledge: 51-59.

Epstein C. (2013). Theorizing agency in Hobbes's wake: the rational actor, the self, or the speaking subject? *International Organization* 67(2): 287-316.

European Parliament (2006) Alleged secret detentions and unlawful inter-state transfers of detainees involving Council of Europe member states. Available at: https://assembly.coe.int/CommitteeDocs/2006/20060606_Ejdoc162006PartII-FINAL.pdf (accessed 30 March 2021)

Fearon JD (1995) Rationalist explanations for war. *International Organization* 49(3): 379-414.

Fearon J and Wendt A (2002) Rationalism v. constructivism: a skeptical view. In: Risse T, Simmons BA and Carlsnaes W (eds) *Handbook of international Relations*. London: Sage Publications, pp. 52-72.

Finnemore M and Sikkink K (1998) International Norm Dynamics and Political Change. *International Organization* 52(4): 887-917.

Forsythe D (1991) *The Internationalization of Human Rights*. Lexington, MA: Lexington Books.

Forsythe DP and McMahon, PC (2016) *American Exceptionalism Reconsidered: US Foreign Policy, Human Rights, and World Order*. London: Routledge.

Foreign Affairs Committee (2011) Third Report from the Foreign Affairs Committee Session 2012-13, the FCO'S Human Rights Work in 2011. Available at: https://www.gov.uk/government/publications/the-foreign-offices-human-rights-work-in-2011 (accessed 31 March 2021).

Frieden JA (1999) Actors and preferences in international relations. In Lake D and Powell R (eds) *Strategic choice and International Relations*. Princeton: Princeton University Press, pp.39-76.

Geddes B (2003) *Paradigms and sand castles: Theory building and research design in comparative politics*. Ann Arbor: University of Michigan Press.

Gerring J (2004) What is a case study and what is it good for? *American Political Science Review* 98(2): 341-354.

Gibney M and Skogly S (2012) *Universal human rights and extraterritorial obligations*. Philadelphia: University of Pennsylvania Press.

Gillis JR (1994) *Commemorations: The politics of national identity*. Princeton: Princeton University Press.

Goodman R and Jinks D (2013) *Socializing states: Promoting human rights through international law*. Oxford: Oxford University Press.

Goldsmith JL and Posner EA (2005) *The limits of international law*. Oxford: Oxford University Press.

Grobe C (2010) The power of words: Argumentative persuasion in international negotiations. *European Journal of International Relations* 16(1): 5-29.

Haas P (2004) When does power listen to truth? A constructivist approach to the policy process. *Journal of European Public Policy* 11(4): 569–92.

Hafner-Burton EM (2005) Trading human rights: How preferential trade agreements influence government repression. *International Organization* 59(3): 593-629.

Hafner-Burton EM and Tsutsui, K (2005) Human rights in a globalizing world: The paradox of empty promises. *American Journal of Sociology* 110(5): 1373-1411.

Hafner-Burton EM (2008) Sticks and stones: Naming and shaming the human rights enforcement problem. *International Organization* 62(4): 689-716.

Hafner-Burton EM, Haggard S, Lake DA and Victor DG (2017) The behavioral revolution and international relations. *International Organization* 71(1): 1-31.

Herweg N, Huß C and Zohlnhöfer R (2015) Straightening the three streams: Theorising extensions of the multiple streams framework. *European Journal of Political Research* 54(3): 435-449.

Herweg N, Zahariadis N and Zohlnhöfer R (2018) The multiple streams framework: Foundations, refinements, and empirical applications. *Theories of the Policy Process*: 17-53.

Heupel M. (2020) Indirect accountability for extraterritorial human rights violations. *International Studies Perspectives* 21(2): 172-197.

Heupel M, Hirschmann G and Zürn M (2018) International organisations and human rights: What direct authority needs for its legitimation. *Review of International Studies* 44(2): 343-366.

HM Government (2019) The Principles relating to the detention and interviewing of detainees overseas and the passing and receipt of intelligence relating to detainees. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/818306/20190718_The_Principles_relating_to_the_detention_and_interviewing_of_detainees_overseas.pdf (accessed 28 March 28 2021).

Hollyer JR and Rosendorff BP (2011) Do human rights agreements prolong the tenure of autocratic ratifiers. *New York University Journal of International Law and Politics* 44(1): 791-811.

Horowitz J and Cammarano S (2013) 20 Extraordinary Facts about CIA Extraordinary Rendition and Secret Detention. Available at: https://www.justiceinitiative.org/voices/20-extraordinary-facts-about-cia-extraordinary-rendition-and-secret-detention (accessed 13 September 2021).

Human Rights Watch (2006) Undermining the Torture Ban. Available at: https://www.hrw.org/legacy/backgrounder/eca/uk1106/4.htm (accessed 16 December 2021).

Ikenberry GJ (2019) *After victory*. Princeton: Princeton University Press.

IPCO (2021) Consultations. Available at: https://www.ipco.org.uk/publications/consultations/consultations-documents/page/2/ (accessed 30 March 2021).

Intelligence and Security Committee of Parliament (ISC) (2018a) Detainee Mistreatment and Rendition_ 2001-2010. Available at: https://fas.org/irp/world/uk/isc-detainee.pdf (accessed 30 March 2021).

Intelligence and Security Committee of Parliament (ISC) (2018b) Detainee Mistreatment and Rendition_ Current Issues. Available at: https://fas.org/irp/world/uk/isc-detainee2.pdf (accessed 30 March 2021).

Jones MD (2003) Bounded rationality and political science: Lessons from public administration and public policy. *Journal of Public Administration Research and Theory* 13(4): 395-412.

Kalin W (2007) Extraterritorial Applicability to the Convention against Torture. *The City University of New York Law Review* 11(2): 293-301.

Karp DJ (2014) *Responsibility for human rights: Transnational corporations in imperfect states*. Cambridge: Cambridge University Press.

Keck ME and Sikkink K (1998) *Activists beyond borders: Advocacy networks in international politics*. Ithaca: Cornell University Press.

King N, Horrock C and Brooks J (2018). *Interviews in qualitative research*. London: Sage Publications.

Krain M (2012) J'accuse! Does naming and shaming perpetrators reduce the severity of genocides or politicides? *International Studies Quarterly* 56(3): 574-589.

Lake DA (2011) *Hierarchy in International Relations*. Ithaca: Cornell University Press.

Jackson R (2005) *Writing the War on Terrorism*. Manchester: Manchester University Press.

Jackson R (2015) The Epistemological Crisis of Counterterrorism. *Critical Studies on Terrorism* 8(1): 33–54.

Jacobs MD (2017) A 'Paradigm of Prevention': United States Counterterrorism Strategies in a Transnational World. In Romaniuk SC, Grice F, Irrera D and Webb S (eds) *The Palgrave Handbook of Global Counterterrorism Policy*. London: Palgrave Macmillan, pp. 163-194.

Jupille J, Caporaso JA and Checkel JT (2003) Integrating institutions: Rationalism, constructivism, and the study of the European Union. *Comparative Political Studies* 36(1): 7-40.

Keck ME and Sikkink K (1999) Transnational advocacy networks in international and regional politics. *International Social Science Journal* 51(159): 89-101.

Kingdon JW (1984) *Agendas, alternatives, and public policies*. Boston: Little, Brown and Company.

Koh HH (2018) Trump v. International Law: Who's winning? *Australian International Law Journal* 24(1): 1-13.

Leech BL (2002) Asking questions: Techniques for semi-structured interviews. *PS: Political Science & Politics* 35(4): 665-668.

Lin LW (2010) Corporate social responsibility in China: Window dressing or structural change. *Berkeley Journal of International Law* 28(1): 64-100.

Luban D (2007) Liberalism, torture, and the ticking bomb. In Lee SP (ed) *Intervention, Terrorism, and Torture*. Dordrecht: Springer, pp. 249-262.

Mahoney J and Goertz G (2006) A tale of two cultures: Contrasting quantitative and qualitative research. *Political Analysis* 14(3): 227-249.

Manners I and Whitman R (2003) The 'difference engine': constructing and representing the international identity of the European Union. *Journal of European Public Policy* 10(3): 380-384.

May T (2016) *Keynote speech at Tory conference in full*. Available at: https://www.independent.co.uk/news/uk/politics/theresa-may-speech-tory-conference-2016-full-transcript-a7346171.html (accessed 5 May 2021).

Miller T, Birch M, Mauthner M and Jessop, J (2012) *Ethics in qualitative research*. London: Sage Publication.

Mir A (2018) What explains counterterrorism effectiveness? Evidence from the US drone war in Pakistan. *International Security* 43(2): 45-83.

Mitchell NJ (2012) *Democracy's Blameless Leaders*. New York: New York University Press.

Mor BD (2009) Accounts and impression management in public diplomacy: Israeli justification of force during the 2006 Lebanon war. *Global Change, Peace & Security* 21(2): 219-239.

Morrow J (1999) The Strategic Setting of Choices: Signaling, Commitment and Negotiation. In Lake DA and Powell R (eds) *Strategic Choice and International Relations*. Princeton: Princeton University Press, pp.78-114.

Murdie AM and Davis DR (2012) Shaming and blaming: Using events data to assess the impact of human rights INGOs. *International Studies Quarterly* 56(1): 1-16.

Neumayer E (2005) Do international human rights treaties improve respect for human rights? *Journal of Conflict Resolution* 49(6): 925-953.

Nowak M (2010) Obligations of States to Prevent and Prohibit Torture in an Extraterritorial Perspective. In: Gibney M and Skogly S (eds) *Universal Human Rights and Extraterritorial Obligations*. Philadelphia: University of Pennsylvania Press.

Office of Legal Counsel (2002) Memorandum for John Rizzo Acting General Counsel of the Central Intelligence Agency - Interrogation of al Qaeda Operative. Available at: https://www.justice.gov/olc/file/886296/download (accessed 27 July 2021).

Peksen D (2012) Does foreign military intervention help human rights? *Political Research Quarterly* 65(3): 558-571.

Piccone T (2018) China's long game on human rights at the United Nations. Available at: https://www.brookings.edu/research/chinas-long-game-on-human-rights-at-the-united-nations/ (accessed 01 September 2021).

Poe SC (2004) The Decision to Repress: An Integrative Theoretical Approach to Research on Human Rights and Repression. In: Carey SC and Poe SC (eds) *Understanding Human Rights Violations: New Systematic Studies*. Aldershot: Ashgate Publishing, pp. 16–38.

Raunio T and Wagner W (2020) The Party Politics of Foreign and Security Policy. *Foreign Policy Analysis* 16(4): 515-531.

Reinold T and Zürn M (2014) "Rules about Rules" and the Endogenous Dynamics of International Law: Dissonance Reduction as a Mechanism of Secondary Rule-making. *Global Constitutionalism* 3(2): 236-273.

Risse T, Risse-Kappen T, Ropp SC and Sikkink K (1999) *The power of human rights: International norms and domestic change*. Cambridge: Cambridge University Press.

Risse T, Ropp SC and Sikkink K (2013) *The persistent power of human rights: From commitment to compliance*. Cambridge: Cambridge University Press.

Reus-Smit C (2013). Constructivism. In: Burchill S., Linklater A, Devetak R, Donnelly J, Paterson M, Reus-Smit C and True J (eds) *Theories of international relations*. New York: Palgrave Macmillan, pp. 188-212.

Ryan L (2012) Cosmopolitan memory and national memory conflicts: On the dynamics of their interaction. *Journal of Sociology* 50(4): 501–514.

Sanders R (2018) Human rights abuses at the limits of the law: Legal instabilities and vulnerabilities in the 'Global War on Terror'. *Review of International Studies* 44(1): 2-23.

Scharpf FW (2018) *Games real actors play: Actor-centered institutionalism in policy research*. London: Routledge.

Schimmelfennig F (2001) The community trap: Liberal norms, rhetorical action, and the eastern enlargement of the European Union. *International Organization* 55(1): 47-80.

Schimmelfennig F (2005) Strategic calculation and international socialization: Membership incentives, party constellations, and sustained compliance in Central and Eastern Europe. *International Organization* 59(4): 827-860.

Schimmelfennig F and Sedelmeier U (2004) Governance by conditionality: EU rule transfer to the candidate countries of Central and Eastern Europe. *Journal of European Public Policy* 11(4): 661-679.

Senate Select Committee on Intelligence (2014) Committee Study of the Central Intelligence Agency's Detention and Interrogation Program. Available at: https://www.feinstein.senate.gov/public/_cache/files/7/c/7c85429a-ec38-4bb5-968f-289799bf6d0e/D87288C34A6D9FF736F9459ABCF83210.sscistudy1.pdf (accessed 15 February 2021).

Shannon BN, McGee ZA and Jones BD (2019) Bounded rationality and cognitive limits in political decision making. Available at: https://oxfordre.com/politics/view/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-961 (accessed 05 July 2021).

Shor E (2008) Conflict, terrorism, and the socialization of human rights norms: The spiral model revisited. *Social Problems* 55(1): 117-138.

Siggelkow N (2007) Persuasion with case studies. *Academy of Management Journal* 50(1): 20-24.

Simmons BA (2013) From Ratification to Compliance: Quantitative Evidence on the Spiral Model. In: Risse T, Ropp SC and Sikkink K (eds) *The Persistent Power of Human Rights: From Commitment to Compliance*. Cambridge: Cambridge University Press, pp. 43-59.

Simon HA (1997) *Models of bounded rationality: Empirically grounded economic reason*. Cambridge: MIT Press.

Snidal D (2002) Rational choice and international relations. In: Carlsnaes W, Risse T and Simmons BA (eds) *Handbook of international Relations*. London: Sage Publication, pp.85-111.

Straus S (2007) Second-Generation Comparative Research on Genocide. *World Politics* 59(3): 476–501.

Trump D (2017) Transcript: ABC News anchor David Muir interviews President Trump, *ABC News*. Available at: https://abcnews.go.com/Politics /transcript-abc-news-anchor-david-muir-interviews-president/story?id=45047602 (accessed 4 May 2021).

US Department of State (2015) Release of the 2015 Country Reports on Human Rights Practices. Available at: https://2009-2017.state.gov/secretary/remarks/2016/04/255799.htm (accessed 17 May, 2021).

Van Evra S (1997) *Guide to Methods for Students of Political Science*. Ithaca, NY: Cornell University Press.

Vreeland JR (2008) Political institutions and human rights: Why dictatorships enter into the United Nations Convention Against Torture. *International Organization* 62(1): 65-101.

Waldner D (2012) Process tracing and causal mechanisms. In: Kincaid H (ed) *The Oxford Handbook of Philosophy of Social Science.* Oxford: Oxford University Press, pp. 65-84.

Weaver RK (1986) The politics of blame avoidance. *Journal of Public Policy* 6(4): 371-398.

Weedon C and Jordan G (2012) Collective memory: theory and politics. *Social Semiotics* 22(2): 143-153.

Youngs R (2015) Exploring" Non-Western Democracy". *Journal of Democracy* 26(4): 140-154.

Zahariadis N (2003) *Ambiguity and choice in public policy: Political decision making in modern democracies*. Washington D.C.: Georgetown University Press.

Zhu X (2008) Strategy of Chinese policy entrepreneurs in the third sector: Challenges of technical infeasibility. *Policy Sciences* 41(4): 315-334.

Zohlnhöfer R (2016) Putting together the pieces of the puzzle: Explaining German labor market reforms with a modified multiple-streams approach. *Policy Studies Journal* 44(1): 83-107.

Zulaika J (2009) *Terrorism. The Self-Fulfilling Prophecy*. Chicago: University of Chicago Press.

Seeing Reason or Seeing Costs?

The United States, Counterterrorism, and the Human Rights of Foreigners

Abstract

It is well known that in the wake of 9/11, the United States committed various extraterritorial human rights violations, that is, human rights violations against foreigners outside of its territory. What is less known is that the United States has gradually introduced safeguards that are, at least on paper, meant to prevent its counter-terrorism policies from causing harm to foreigners abroad or, at least, to mitigate such harm. Based on three case studies on the development of safeguards related to torture, targeted killing, and mass surveillance, we show that two mechanisms, coercion and strategic learning, deployed either on their own or in combination, can account for the development of such safeguards. By contrast, we found no evidence of a third mechanism, moral persuasion, having any direct effect. In other words, US policymakers opt to introduce such safeguards either when they face pressure from other states, courts, or civil society that makes immediate action necessary or when they anticipate that not introducing them will, at a later date, result in prohibitively high costs. We did not find evidence of US policymakers establishing safeguards because they deemed them morally appropriate. From this we conclude that, although the emerging norm that states have extraterritorial (and not just domestic) human rights obligations may not have been internalized by key US policymakers, it nevertheless has a regulative effect on them insofar as the fact that relevant others believe in the norm restricts their leeway and influences their cost–benefit calculation.

Keywords: Deterritorialization, human rights, International Relations, national security, norms, War on Terror

1. Introduction

The United States (US) has committed extraterritorial human rights violations in its response to 9/11. Terrorist suspects have been tortured in Abu Ghraib, arbitrarily detained in Guantánamo Bay, or brought to black sites; terrorist suspects and civilians have lost their lives in targeted killing operations; citizens around the world have had their privacy rights infringed through mass surveillance programs; and refugees have faced ever larger obstacles to entering the US. Nonetheless, the US has also begun to introduce measures that are designed to prevent, or at least mitigate, the harmful effects of its policies on foreigners abroad. Government agencies and Congress have introduced safeguards that, among other things, prohibit torture

and extraordinary rendition, allow Guantánamo inmates to challenge their detention, specify criteria for targeted killing operations, and ban intelligence agencies from indiscriminately spying on foreigners. Few of these safeguards effectively guarantee that extraterritorial human rights violations will not occur; some of them are almost certainly paper tigers. Yet they demonstrate that US policymakers recognize that they cannot treat foreigners beyond their borders simply as they like. The key question remains, however, why has the US established such safeguards and imposed restrictions on itself?

Research on extraterritorial human rights obligations is primarily the domain of legal scholars, a number of whom have criticized a "paradox in international human rights law", namely that human rights were meant to be universal, while international human rights conventions have traditionally been perceived as primarily containing obligations towards a state's own nationals or effective within its own territory (Gibney et al., 1999: 267). Since human rights are essentially protections against the unchecked exercise of power, however, the "territoriality paradigm" (Vandenhole and Van Genugten, 2015:1) of international human rights law no longer suits a world in which states increasingly interact with non-citizens beyond their borders, suggesting that states also have obligations towards the latter (King, 2009: 522). Legal scholars also interpret the jurisprudence emanating from national and international courts and United Nations (UN) treaty bodies and, although the matter is still contested, have identified a growing recognition in relevant case law that states have extraterritorial human rights obligations that arise in contexts of territorial and factual control, though these are largely restricted to negative obligations (Gammeltoft-Hansen and Vested-Hansen, 2017; Skogly, 2006). This scholarship is undoubtedly of great value. Legal scholars, however, have not focused on gathering insights into why states introduce safeguards to minimize the harm their policies cause to foreigners abroad.

Political scientists, on the other hand, are very interested in why states commit to and comply with human rights standards. Some scholars have shown that shaming by transnational non-governmental organizations (NGOs) can improve states' compliance (Murdie and Davis, 2012); some have argued that litigation is an important strategy for holding governments accountable (Simmons, 2009: 129–135); others have demonstrated that states respond to material incentives (Hafner-Burton, 2005); finally, states may undergo a socialization process and be persuaded to believe in the value of human rights norms (Risse et al., 1999). How these or other mechanisms operate when extraterritorial human rights obligations are involved rather than domestic ones has not, however, been investigated. Moreover, scholars who have looked specifically into how

US counterterrorism law has evolved since 9/11 only describe the changes (Setty, 2015, but see Abel 2018); research that explains the emergence of US safeguards for foreigners abroad is based on single case studies (Sikkink, 2013).

In light of this gap, this article investigates why the US has introduced protections for foreigners outside US territory against harm inflicted by its counterterrorism policies. We conceptualize three causal mechanisms – moral persuasion, strategic learning, and coercion – and examine which mechanism(s) account(s) for the emergence of safeguards. We do so on the basis of three case studies concerning the development of safeguards related to the right of detainees not to be tortured, the right to life in targeted killing operations, and the right to privacy in the context of foreign mass surveillance.

We have found two mechanisms, coercion and strategic learning, that are able, on their own or in combination, to explain why the US has introduced safeguards. Although we did not find one mechanism that operated in all cases, our findings suggest that the introduction of safeguards follows a distinct pattern. Most importantly, our findings indicate that cost–benefit calculations by US policymakers were critical, whereas moral persuasion for the most part had no direct effect on them. This does not imply, however, that the norm that states have extraterritorial human rights obligations has not begun to diffuse internationally. Rather, it indicates that the norm primarily exerts a regulative effect. Key US policymakers may not have internalized the norm, but its existence has swayed their cost–benefit analyses in favor of introducing safeguards, suggesting that moral persuasion has an indirect effect.

This article consists of three sections. We first conceptualize our causal mechanisms and introduce the article's research design. Subsequently, we present and interpret the findings of the empirical analysis. The concluding section also outlines avenues for future research.

2. Theory

Mechanisms are chains of events that connect a starting point with an outcome. Following the logic of Coleman's (1986) macro–micro link in social action, we assume that events at the macro level (human rights violations) trigger action at the micro level (interventions by individuals/groups; input processing by policymakers), which, in turn, results in changes at the macro level (safeguards). Our mechanisms therefore all follow the same logic: Extraterritorial human rights violations evoke an intervention; policymakers process the intervention; as a

result, they establish safeguards. Importantly, a mechanism can break down at any link in the chain if the conditions that facilitate the transition between its components are not given (Bennett and Checkel, 2015: 12).

We derive three mechanisms from existing theories on state commitment to and compliance with human rights norms. They are conceptualized as unique inasmuch as their individual components do not overlap (Goertz, 2017: 48). First, policymakers may become convinced of the inherent value of a norm and enact rules that reflect this conviction (moral persuasion). Second, policymakers may, after thorough reflection and in the absence of immediate pressure, realize that it is in their long-term strategic interest to prevent the potential negative consequences of rights violations by establishing safeguards (strategic learning). Third, policymakers may face immediate pressure in the form of material sanctions, shaming, or litigation, to which they feel forced to respond (coercion). Why actors decide to intervene with policymakers is not conceptualized as part of the mechanisms. Hence, when we assess whether a mechanism is present in a specific case, we only consider its direct effect (how it works on policymakers), not its indirect effect (how it works on actors who try to influence policymakers). We do, nonetheless, address such indirect effects when we discuss our findings. Below, we provide a short conceptualization of the mechanisms, the following table provides a summary; for a more detailed conceptualization see sections 1.1 and 1.3 of the appendix.

Starting Point →	Intervention →	Processing →	Outcome
Moral persuasion			
Extraterritorial human rights violations →	Moral arguments →	Acceptance of norm →	Safeguards
Strategic learning			
Extraterritorial human rights violations →	Strategic arguments →	Anticipation of future strategic gains →	Safeguards
Coercion			
Extraterritorial human rights violations →	Material sanctions, shaming, or litigation →	Perception of an urgent need to react →	Safeguards

Table 1: Mechanisms

Moral persuasion

The moral persuasion mechanism, applied to our case, is, like any other, triggered by state actors committing human rights violations on foreign territory. Policymakers in the norm-violating state are then confronted with arguments by norm entrepreneurs as to why their behavior is morally wrong in order to convince them of the intrinsic value of the norm in question (Finnemore and Sikkink, 1998). Norm entrepreneurs may, for instance, argue that, given the universal nature of human rights, it is morally unjustifiable for states to violate the human rights of foreigners outside of their territory. Offending policymakers may, as a result, experience cognitive dissonance because they realize that their behavior is not in line with normative expectations (see Reinold and Zürn, 2014). In this process, they may become convinced of the norm's inherent value and conclude that their behavior has been morally wrong (Risse, 2000). In our case, key actors in government and/or Congress might become convinced that it is morally appropriate to insure that foreigners abroad are not harmed by US counterterrorism policies and, consequently, take steps to comply with this norm by establishing new regulations or laws. While most violations may trigger criticism on moral grounds, the successful processing of moral arguments is key in determining whether moral persuasion was the causal mechanism behind a particular safeguard.

Strategic learning

In the strategic learning mechanism, policymakers in the norm-violating state are confronted with strategic (as opposed to moral) arguments as to why their norm-violating behavior is likely to undermine their own or their country's long-term interests (see Grobe, 2010). Crucially, this happens before immediate pressure to act builds up. Strategic arguments may be put forward by norm entrepreneurs who consider them more convincing than moral ones or by other stakeholders who perceive their own or their national interests to be in danger. Applied to our case, strategic arguments might be made that extraterritorial human rights violations could jeopardize key objectives in the fight against terrorism or undermine the US's authority to demand compliance with human rights norms from other states. Having processed these arguments, policymakers in the perpetrator state might determine that continuing the norm-violating behavior is likely to undermine their strategic goals.[1] Rational actors are expected to weigh the anticipated costs of policy change against the perceived benefits. They will then be likely to conclude that the benefits associated with reforms will be greater than any losses such

reforms might entail (Downs et al., 1996). For example, US policymakers might conclude that they would likely put off their allies in the future or provoke further terrorist violence by pressing ahead without safeguards. If these anticipated costs are deemed to exceed the perceived strategic benefits of their questionable policies, then policymakers will opt for reforms based on a logic of prevention, even if they are not being directly forced to alter their policies.

Coercion

In the coercion mechanism, the intervention that follows norm-violating behavior comes in the form of tangible, immediate pressure and not argument (either moral arguments in the moral persuasion mechanism or instrumental arguments in the strategic learning one). Norm-violating policymakers face immediate negative consequences as a result of their behavior and therefore feel compelled to react. The critical difference from the strategic learning mechanism is that policymakers come to the conclusion that a response is necessary quickly and do not spend time carefully weighing the potential costs and benefits of introducing safeguards. Coercion can come in three variants – material sanctions, shaming, and litigation. They all follow the same basic logic but involve different intervention strategies applied by different actors.

In the material sanctions variant, norm-violating states face material punishment (Donno and Neureiter, 2018; Hafner-Burton, 2005) from other states whose nationals are harmed, international organizations (IOs), or private actors. The US government might, for example, introduce safeguards in response to key allies announcing that they will oust US troops from their territory unless safeguards for their nationals are introduced in line with international human rights law. If policymakers in the target state see the sanctions as an immediate threat to their interests, they will perceive an urgent need to react. US policymakers might consider their allies' response to be prohibitively costly and feel compelled to provide safeguards to induce their allies to lift any sanctions or refrain from implementing them.

In the shaming variant, immediate pressure for policy change may again come from norm entrepreneurs with an honest interest in norm-compliant behavior. This time, however, instead of trying to influence norm-violators with convincing arguments, they tarnish their reputation (Krain, 2012; Murdie and Davis, 2012). Shaming can also be a strategy for actors who do not care about the violated norm but see an opportunity to publicly attribute blame to another actor. Policymakers in the target state perceive reputational damage as a result of a public attribution of blame. US policymakers, for example, might take note of any damage to the US's reputation

as a country that respects human rights and worry that this may have unwanted knock-on effects. Consequently, they introduce safeguards to restore the country's reputation.

In the litigation variant, immediate pressure comes from a court that issues a judgment that identifies a violation of international human rights law and demands safeguards (Duffy, 2018; Simmons, 2009). It might be a domestic court in the target state, a foreign court authorized to exercise universal jurisdiction, or an international court with jurisdiction over the respective state's extraterritorial human rights-related behavior. In order for the litigation mechanism to operate, policymakers must process the intervention and perceive an obligation to implement the court's decision and abide by its interpretation of applicable human rights law. As a direct response to an actual court judgment or to the imminent threat of one, US policymakers would therefore choose to introduce reforms to avoid the negative consequences of non-compliance with a court decision.

3. Research Design

Our research design combines the strengths of process tracing with those of a design based on the analysis of several cases (Bennett and Checkel, 2015: 21). We use deductive process tracing to follow the selected causal mechanisms across three cases. Deductive process tracing requires the conceptualization of one or several hypothetical mechanisms before investigating whether empirical evidence for the operation of the mechanism(s) can be identified in any of the cases, working on the basis that a mechanism has explanatory value only if all of its components are present (Beach and Pedersen, 2019: 255, 260).

We do not expect that all cases in which the US establishes safeguards for foreigners abroad will follow the same path. Instead, we are open to the possibility of equifinality and therefore in each case trace different mechanisms that might plausibly be expected to account for the establishment of any safeguards (Schimmelfennig, 2015: 106-7).[2] Similarly, we assume that a single mechanism may not always be able to do justice to the complexity of policy change. Hence, we are also open to the possibility that the interplay of multiple mechanisms may account for the outcome in a specific case (see sections 1.2 and 1.3 of the appendix). Following a sequential pattern, one mechanism may trigger another; alternatively, following a cumulative logic, two mechanisms may operate simultaneously, together leading to the outcome (Beach and Rohlfing, 2018). However, to avoid overdetermined "kitchen sink arguments in which

everything matters" (Checkel, 2006: 367), it is necessary to distinguish between genuinely causal and spurious relationships and to take into account the persuasiveness of the empirical evidence found for each mechanism (Bennett and George, 2005: 222). Finally, it should not be overlooked that when mechanisms interact, they do not necessarily reinforce each other but may crowd one another out instead if, for instance, their logics of influence are incompatible (Goodman and Jinks, 2013).

We examine the explanatory value of the selected mechanisms in three parallel case studies in which the US, having committed extraterritorial human rights violations following 9/11, established what we call 'extraterritorial human rights safeguards'. Specifically, we consider cases where the US's negative obligations not to violate the right to be free from torture (in the context of detainee treatment), the right to life (in targeted killing operations), and the right to privacy (in relation to foreign mass surveillance) were involved. Empirical evidence has been gathered from primary and secondary sources and from 43 interviews with US policymakers, their staff, bureaucrats and experts conducted in 2017 and 2019.

Case selection was based on two considerations. First, in line with the logic of deductive process tracing, we chose cases in which both the mechanisms' common starting point (extraterritorial human rights violations) and common endpoint (the introduction or enhancement of safeguards designed to prevent or mitigate harm to foreigners outside the US) were given (Beach and Pedersen, 2019: 98-99, 258). We thus follow Goertz's (2017: 59) advice to begin with cases in which hypothesized mechanisms can be explored in their entirety, before moving on to others that lack either the mechanisms' starting point or their endpoint. Indicators that existing provisions are being enhanced are increases in the scope of the rights violations they cover, the degree of obligation they imply, their precision, the number of their beneficiaries, or the addition of complaint provisions to preventive measures (see Heupel and Hirschmann, 2017). Second, by conducting three case studies with the purpose of "accumulating systematically within-case causal inferences" (Goertz, 2017: 173), we take the expectation of equifinality seriously, because conducting several case studies enables us to detect different causal mechanisms (Hall, 2013: 28). Moreover, as our cases differ in important ways - for instance with respect to the type of right violated or the actors who suffer rights violations - we can probe the mechanisms in different contexts and learn something about the conditions under which they occur.

To present the results of small-N process tracing in a journal article, one has to reconcile the contradictory demands of achieving internal and external validity. Given the space limits, it is impossible to give an account of all the important empirical information. Nonetheless, we have applied the following strategies to achieve a balance: The analysis that forms the basis of our brief case summaries has been as detail-oriented as process tracing demands, but the case summaries presented concentrate solely on key events and background conditions. At the same time, they mirror the structure provided in the way that the mechanisms are conceptualized, and we provide as many references as possible, again respecting the space limitations given. Finally, we concentrate on the mechanisms that operated as conceptualized (coercion and strategic learning) and restrict discussion of the role of moral persuasion to section 5 of this article and section 2 of the appendix.

4. Empirical Analysis

Our empirical analysis confirms our expectation of equifinality. We found two mechanisms, coercion and strategic learning, that can, either on their own or in combination (with coercion preceding strategic learning), explain the emergence of safeguards that provide protections for foreigners outside the US against harm caused by US counterterrorism policies (see Table 2). Moral persuasion did not have a direct effect on US policymakers but, as we will discuss in the next section and in section 2.2 of the appendix, had important indirect effects. The remainder of this section provides short summaries of our three cases.

Human Rights	Operational Context	Casual Mechanism(s)
Right not to be tortured	Detainee treatment	Coercion + Strategic learning
Right to life	Targeted killing	Strategic learning
Right to privacy	Mass surveillance	Coercion

Table 2: Results

Coercion + Strategic Learning: Detainee Treatment and the Right not to be Tortured

The development of extraterritorial anti-torture safeguards arises out of a sequential combination of two mechanisms, an initial phase of coercion being followed by a phase of strategic learning. In the aftermath of 9/11, the perceived need for better intelligence opened the door to interrogation techniques for foreign terror suspects that even President Barack Obama, when commenting on practices endorsed by the Bush administration, later described as a violation of the right to be free from torture (The White House, 2014b). In 2002 the Office of Legal Counsel issued memoranda (later known as the 'torture memos') which introduced so-called Enhanced Interrogation Techniques (EITs). They also outlined why the protections of the Geneva Conventions were not applicable to Al-Qaida and how the Taliban could be denied prisoner-of-war status (Office of Legal Counsel, 2002b). According to these memos, the ten EITs, including waterboarding, walling, and sleep deprivation, would not result in long-lasting mental harm, nor did they constitute torture or cruel, inhuman, or degrading treatment (CIDT), because they were not "equivalent in intensity to the pain accompanying serious physical injury, such as organ failure, impairment of bodily function, or even death" (Office of Legal Counsel, 2002a: 1). The actual application of the techniques, however, stood in stark contrast to the Counsel's legal opinion. Official government reports confirm, for instance, that one detainee had been waterboarded 183 times within 14 days, while another died of hypothermia after he was held partially nude at low temperatures for at least 48 hours (CIA Inspector General, 2004: 74).

Alberto Mora, General Counsel of the Navy and a firm opponent of torture, tried, upon learning of the use of torture in the Guantánamo Bay detention facility in late 2002, to convince the Department of Defense to ban any form of torture, describing torture as a betrayal of American values (Department of the Navy, 2004). His efforts did not lead to any significant reforms, however (Mayer, 2006). Meanwhile, NGOs began to publicly denounce the US government's interrogation practices (American Civil Liberties Union, 2019). Due to the top-secret classification of the interrogation program, it was difficult to present evidence to back the accusations made, so that the topic's 'campaign-ability' initially remained low.[3] Similarly, early media reports alluded to abuses in US-run prisons in Iraq, but because they lacked photographic evidence, were received with little interest by the public (Hanley, 2003).

This, however, changed with the Abu Ghraib leaks in spring 2004, when a series of very graphic photos depicting severe cases of torture and CIDT were broadcast nationwide on CBS News.[4] Subsequent leaks of the 'torture memos' and abusive detainee treatment in Guantánamo Bay

further boosted the public uproar,[5] allowing NGOs to initiate national and international shaming campaigns, that increasingly framed detainee abuse as a war crime (Amnesty International, 2004). At the same time, growing media coverage spread the story globally, putting policymakers under immediate pressure to act (Abel, 2018: 45-105; Hersh, 2004).

State Department officials quickly became aware of the reputational damage the scandal had produced.[6] Moreover, Senator John McCain, a long-time opponent of torture, talked to key stakeholders in Congress, arguing that failure to restore the US's reputation would put US military personnel into direct danger.[7] Although McCain and others also made moral arguments, pointing to commitment to the torture ban being part of the US identity, such rhetoric had little impact.[8] Finally, in autumn 2005, in direct response to the shaming campaign triggered by the Abu Ghraib leaks, the Senate passed the Detainee Treatment Act (DTA) as amendment 1977 to the Defense Appropriations Bill for 2005 by 90 votes to 9 (US Congress 2005). The new law bound the military to the interrogation techniques listed in the Army Field Manual (AFM) (US Congress 2005: Sec. 1002), making EITs an illegal practice for the Department of Defense. The DTA had its shortcomings, as it did not impose the same obligations on the Central Intelligence Agency (CIA) and introduced a good faith assumption for prior actions by US military personnel (US Congress 2005: Sec. 1004). Nevertheless, it constituted an important reform as it prohibited, thenceforth, any use of torture and CIDT by the US military at all times and without geographical limitations, introducing a significant change to previous policy standards that had only applied such safeguards to detention facilities under officially recognized US jurisdiction.

NGOs, nevertheless, remained concerned about the absence of regulations for the CIA and criticized the lack of transparency surrounding the implementation of the new law.[9] McCain, similarly, fiercely resented the loopholes in the legislation[10] that enabled President Bush to declare that he would interpret the bill "in a manner consistent with the constitutional authority of the President" (The White House, 2005). The NGOs therefore decided to push for further reforms. Yet, with the US public growing increasingly indifferent to the issue, activists became aware of the declining potential of shaming. They felt forced to change their strategy and instead turned to feeding strategic arguments to key policymakers about the anticipated negative consequences for the US of using torture.[11] Human Rights First, one of the key NGOs working on the issue, began to cooperate with retired military and CIA interrogators to convince Congress and the 2008 presidential candidates of the need to abolish any form of torture for all agencies. The retired interrogators argued that torture was "ineffective, unlawful and

counterproductive" and warned that its use "facilitates enemy recruitment, misdirects or wastes scarce resources, and deprives the United States of the standing to demand humane treatment of captured Americans" (Human Rights First, 2008). Presidential candidate Obama engaged actively in such debates, trying to sharpen his understanding of the problem and develop policy proposals with the experts.[12] Eventually, on his second day in office surrounded by the same retired generals and flag officers who had previously advised him about the expected costs of inactivity, President Obama signed Executive Order (EO) 13491 Ensuring Lawful Interrogations that referenced the Convention Against Torture (CAT) and made compliance with the AFM mandatory for all government agencies, including the CIA (The White House, 2009).

EO 13491 did not have the authority of a law, however. After the release of the Senate Select Committee's Study of the Central Intelligence Agency's Detention and Interrogation Program in 2014, the issue was back on the agenda. Once more, experts and scholars publicly spoke out against torture, arguing that not only was it futile, it constituted a looming threat to national security and the safety of members of the US military.[13] Against this background, Senators McCain and Feinstein introduced an Amendment to the National Defense Authorization Act of 2016. In the subsequent debate, many Republicans opposed creating any sort of accountability for past actions, even though they generally opposed torture.[14] Additionally, concerns about reducing the CIA's flexibility by binding it to the AFM coincided with open criticism of the AFM itself, which had originally been designed not as a general interrogation standard but as basic guidance for regular soldiers.[15] Nonetheless, the new law would regulate governmental agencies' conduct going forward and thus not only prevent future reputational damage, but also reduce national security risks deriving from unlawful interrogation techniques. In addition, writing EO 13491 into law would forestall arbitrary changes by future presidents (Senate Select Committee on Intelligence, 2014: vii). After having considered both its anticipated costs and benefits, Congress ultimately passed the amendment in December 2015. The new law made compliance with the AFM mandatory for all government agencies in any armed conflict, enshrining mandatory compliance with the Geneva Conventions and other core points of EO 13491. Likewise, the bill introduced a report on best practices and a regular review of the AFM to ensure the US's compliance with its international obligations (US Congress, 2015: Sec. 1045).

In summary, the development of extraterritorial anti-torture safeguards can be traced back to coercive pressure in the shape of shaming, followed by strategic learning about the potentially

negative consequences of inaction. Although US policymakers were also confronted with moral arguments against torture and there were public officials who abhorred torture as a matter of principle, there is no evidence that safeguards would have been introduced without pressure and strategic considerations. The safeguards are strong in that they clearly outlaw torture for all government agencies and have survived the Trump presidency despite President Trump's opposition to such limitations. Due to the secrecy surrounding CIA operations, however, hidden cases of torture or other forms of mistreatment in US detention facilities outside of US territory may still exist.

Strategic Learning only: Targeted Killing and the Right to Life

The introduction of extraterritorial safeguards in the US targeted killing program can be traced back to strategic learning within the Obama administration. The use of lethal action[16] rose exponentially from 50 strikes under the Bush administration to 586 strikes under Obama (New America, 2019), violating the right to life of many. Specifically, experts in the field claimed that the practice of targeting terrorist suspects in so-called signature strikes giving them no opportunity to defend themselves legally amounted to extrajudicial killing (Davis et al., 2016: 9), while the steady increase in civilian casualties was criticized as disproportionate and excessive (European Center for Constitutional and Human Rights, 2019). The drone program in non-active combat zones was especially controversial. Although the exact numbers of fatalities for which the program was responsible differ depending on the reporting source, the average estimate indicates a total between 3,400 and nearly 5,000 for the period January 2009 to January 2017 (The Bureau of Investigative Journalism, 2019).

These high fatality numbers and the initial absence of any formal rules regarding the determination of targets motivated various attempts at forcing the US government to consider a change of policy. These met with no success, however. Activists tried to develop a shaming campaign, using their own counts of civilian casualties (New America, 2019) and casting doubt on the US's credibility as a defender of human rights (Center for Civilians in Conflict, 2013; Lawrence, 2013). Yet they failed to change domestic public opinion. By 2013, support for drone strikes abroad among Americans was still at 65%; only 28% opposed them strictly (Brown and Newport, 2013). Similarly, as the US government was able to refer to the principle of self-defense (The White House, 2013b), no litigation made it before a US court, while attempts by

the Pakistani and Afghan governments to sanction the US government proved futile because of their inconsistency.[17]

In the end, however, though coercive pressure did not succeed in triggering policy change, strategic argumentation did. Executive staffers began to worry that the president's quasi-unlimited discretionary power in targeted killing operations could have long-term negative consequences.[18] Likewise, arguments that terrorists might use the US drone program for propaganda or that the US should not miss an opportunity to set standards for other countries that also acquired drone technology played an important role in the strategic debate.[19] Moral arguments against targeted killing with drones were brought forward, too, particularly by civil society actors and UN human rights bodies (e.g. UN Human Rights Council 2010). Yet, there is little evidence that they had a decisive impact on the decision in favor of safeguards.[20]

High-ranking White House staffers and the government's Counterterrorism Working Group engaged with the strategic arguments and ultimately acknowledged that the potential benefits of introducing safeguards would outweigh the anticipated costs associated with them.[21] Although additional bureaucratic procedures would require resources and slow down operational decisions, a public guideline could terminate legal debates and address concerns regarding the absence of rules.[22] A new cross-checking provision for targeting procedures that would involve additional agencies could likewise diminish Pakistan's and Afghanistan's resentment by relying on more intelligence so as to increase strike accuracy and limit future civilian casualties (The White House, 2013b). While the latter would also offset the worries of NGOs, further provisions such as a requirement of a high certainty of minimum civilian casualties would additionally reduce the risk of future shaming campaigns and their negative consequences.[23]

As a result, and in the absence of immediate pressure, Obama signed Presidential Policy Guidance (PPG) – Procedures for Approving Direct Action against Terrorist Targets Located outside the United States and Areas of Active Hostilities (The White House, 2013a) in May 2013, establishing the first official guideline on targeted killing. The document remained rather weak, as it did not restrict the drone program to combat zones, nor significantly improve the policy's transparency. However, it signaled that, "the United States prioritizes, as a matter of policy, the capture of terrorists as a preferred option over lethal action," and that "lethal action should not be proposed or pursued as [a] punitive step or as a substitute for prosecuting a terrorist" (The White House, 2013a: preamble). Additionally, the PPG created, at least on paper,

a rigorous framework for target nomination, interagency review and authorization procedures. Accordingly, any operational plan had to be approved by the operating agency's general counsel(s), the National Security Staff and the Deputies and Principals Committees of the National Security Council (NSC), before being handed to the President for the final decision (The White House, 2013a: Sec.1B). Moreover, the PPG introduced a mandatory requirement of near certainty of no civilian harm, and made an assessment of the compliance of each strike with international law as well as after-action reports and Congressional notifications obligatory (The White House, 2013a: Sec.6). So, the PPG attempted to offer a response to almost all concerns brought forward in the previous debate, testifying to the processing of the strategic arguments presented by internal stakeholders.

Nonetheless, after the enactment of the PPG, the continued elevated use of lethal action remained a matter of concern, particularly as the number of civilian casualties remained high (Greenfield and Hausheer, 2014: 1). State Department officials insistently cautioned that civilian casualties were a looming threat to national security: the high death toll could facilitate terrorist propaganda, creating more radicalized enemies than the strikes were actually eliminating.[24] Additionally, advisors within the White House and the Counterterrorism Working Group grew worried that a continuing lack of transparency and accountability could further increase the risk of reputational damage both domestically and internationally.[25] Moreover, human rights watchdogs continued to voice moral concerns about civilian casualties (Human Rights Watch 2013).

In response, the Counterterrorism Working Group, the NSC, and Department of State officials resumed their discussions, again focusing primarily on strategic arguments. They exchanged views on concerns that reports detailing civilian casualties could reveal secret operations, while creating additional expense. On the other hand, a new policy on civilian casualty mitigation could increase the program's effectiveness tactically and strategically, given that safeguards would improve target accuracy while decreasing the risk of creating new enemies.[26] Likewise, official government reports on civilian casualties could create a greater sense of accountability, which would bolster the US's credibility as a human rights defender.[27] Moral considerations, however, even though they were shared by a number of internal stakeholders, did not take center stage in the debates.[28] Despite some internal opposition, in July 2016 President Obama eventually signed EO 13732 – United States Policy on Pre- and Post-Strike Measures To Address Civilian Casualties in U.S. Operations Involving the Use of Force, which, despite lacking precision throughout, clearly went beyond the PPG. The EO established enhanced

training programs for civilian protection, mandatory periodic performance consultations with the NSC, and regular exchanges on best practice with international partners. It also provided for the development of more precise weapon systems and field intelligence so as to further decrease the number of civilian casualties, while the agencies involved were instructed to take responsibility for any deaths and to publish them, in cooperation with NGOs, in annual reports (The White House, 2016).

In summary, the case is an example of safeguards being introduced as the result of strategic learning. Both the PPG and EO 13732 were promulgated in response to strategic arguments brought forward, in this case, mostly by actors from the executive branch. Both safeguards were written in vague language and, given their nature as executive directives, could be easily altered by future administrations – a feature exploited by President Trump when he revoked parts of the directives in 2019 (The White House, 2019). Nonetheless, the safeguards constitute a small but important step toward the development of safeguards for foreigners in the US targeted killing program.

Coercion only: Foreign Mass Surveillance and the Right to Privacy

In the case of the violation of privacy rights in foreign mass surveillance, safeguards emerged via two iterations of the coercion mechanism.[29] In the aftermath of 9/11, the US government used its existing legal competences (The White House, 1981) and competences newly acquired from Congress (US Congress, 2001, 2008) to expand its capacity for indiscriminate digital surveillance of foreigners. Although the National Security Agency (NSA) receives some of its data on foreigners from foreign governments, its direct access to data on foreign users via US companies is key to its surveillance efforts. It has supposedly hacked the servers of US technology and telecommunications companies and relied on the Foreign Intelligence Surveillance Court to order technology companies like Facebook and Google to turn over user data. Moreover, it has remunerated telecommunication companies for tapping emails that transit their internet cables (Greenwald, 2014). The US government thus treated private companies as "tools of national intelligence" (Farrell and Newman, 2016: 131), benefitting from their collection of huge amounts of data from their foreign customers for its own purposes.

In mid-2013, Edward Snowden, an NSA contractor turned whistleblower, exposed the extent and indiscriminateness of US surveillance practices, feeding confidential information on the doings of the NSA to the Guardian and the Washington Post. The issue of domestic surveillance

briefly dominated the news media in the US, and human rights NGOs and UN agencies publicly blamed the US government for indiscriminate mass surveillance (Amnesty International, 2013; UN Human Rights Committee, 2014). Nonetheless, a public campaign on privacy rights violations of foreigners outside of the US never gained momentum. Not only was the issue overshadowed by the focus on domestic transgressions, it was also inherently difficult to visualize. Consequently, the government did not suffer the kind of reputational harm that would have necessitated immediate action. Moreover, moral arguments as to why safeguards for foreign citizens were appropriate (e.g. Roth 2013) did not motivate policymakers to establish them (see Bignami and Resta, 2018: 364, 378) and, while foreign governments did express public criticism, they were not very vocal as they depended on the US for intelligence sharing (Abel, 2018: 395-473).

Although the US government was not vulnerable to shaming, US technology companies were (Farrell and Newman, 2016: 125, 128). Companies like Facebook and Google depend on the trust of their users. This weak spot was exploited by actors who believed they could use the companies as an instrument to make themselves heard.[30] Accordingly, journalists zoomed in on the supporting role of US technology companies in US foreign surveillance (Greenwald and MacAskill, 2013). Importantly, these accusations not only entailed reputational damage to technology companies but also financial losses when foreign customers began to turn to the services of non-US competitors. Moreover, there was a general assumption that there would be greater losses of market share in the future (Donohue, 2015).

Unwilling to bear the costs of NSA foreign surveillance, technology companies lobbied the US government for privacy safeguards. They made clear that it was extremely important for them that the government sent out a public signal to the effect that it took the issue of privacy safeguards for foreigners abroad seriously (Levy, 2014; Schneier, 2015: 122). They also communicated unequivocally that they were ready to penalize the US government immediately if it did not act. Specifically, they threatened to invest in stronger encryption and to store foreign user data abroad. Technology companies even began to strengthen their encryption capacities as a direct response to the Snowden leaks (Timberg, 2013).

The US government wanted continued access to the tech companies' data from foreign users, but was concerned about the proliferation of data encryption initiatives and believed that the companies would follow through with at least some of their threats. The Obama administration therefore decided quite quickly to send out the public signal the tech companies wanted, and in

early 2014 released Presidential Policy Directive 28 – Signals Intelligence Activities (PPD-28), which announced that the "United States will … impose new limits on its use of signal intelligence in bulk … intended to protect the privacy and civil liberties of all persons, whatever their nationality and regardless of where they might reside" (The White House, 2014a: Sec. 2). Although PPD-28 contains rather vague language, it still constitutes an act of "unprecedented self-limitation"[31] as it marks the first time that the US recognized that foreigners living outside its borders are entitled to privacy protections. It is also striking that the limits on the purposes for which foreigners' data could be used were more precise than those set out in comparable documents issued by most countries of the European Union (EU) (Brown et al., 2015: 3, 19). PPD-28 was also a response to requests by other governments.[32] Nonetheless, it was the intervention of US technology companies, especially the pressure and threat of sanctions from them, that was decisive (see Rascoff, 2016: 662, 669, 688–689), as the US government heavily depended on access to their data specifically and good relations with them more generally (see Donohue, 2015: 35-6).

PPD-28 did not, however, end the pressure on tech companies that collected data from foreign customers and stored it on servers in the US. As public attention abated, new pressure came from the Court of Justice of the European Union (CJEU), which had to decide whether the data from foreign users that Facebook's Irish subsidiary transferred to the US were sufficiently protected against government surveillance. In October 2015 the CJEU issued a landmark judgment that invalidated the Safe Harbor Agreement between the European Commission and the US government that had been used by about 4,500 US companies to transfer EU citizens' personal data to the US. Specifically, the court argued these data were insufficiently protected against NSA access, which made their transfer to the US unacceptable (Court of Justice of the EU, 2015). The affected companies were alarmed and demanded more far-reaching privacy safeguards against foreign surveillance from the US government to satisfy the CJEU and, hence, the EU.[33] The US government feared, again, that inaction would prompt the companies to complicate government access to their data.[34]

Finally, less than four months after the CJEU's judgment, the US Department of Commerce issued the Privacy Shield Framework Principles, a self-certification regime for US companies that transfer data from EU citizens to the US (US Department of Commerce, 2016). The European Commission accepted the Principles as providing adequate protection and allowed the transfer of EU data to the US by companies that self-certified under the new regime. On paper, the Privacy Shield went beyond the Safe Harbor Agreement, as it contained transparency

requirements and provided for an Ombudsperson in the State Department to handle complaints by EU citizens. Critics cast doubt on the effectiveness of the Privacy Shield Principles, however, pointing to the hurdles built into the complaints process and the unreproducible assurances by the US government that the agreement was based on.[35] In the end, the CJEU, in a further landmark judgment of July 2020, invalidated the decision attesting that the Privacy Shield provided adequate protection (Court of Justice of the EU, 2020), once again prompting talks between US and EU regulators on how to establish safeguards that would satisfy the court.

In summary, the small steps the US government has thus far taken to introduce privacy safeguards for its foreign surveillance operations can be traced back to coercion. There is no evidence that policymakers would have been persuaded of the inherent value of the right to privacy as it applies to foreigners residing outside of the US or that they would have carefully weighed the potential future costs and benefits of introducing privacy safeguards for foreigners.[36] Rather, safeguards were introduced as an immediate response to sanctions, or the threat of them, by US technology companies.

5. Interpretation

Based on our analysis, there appears to be a discernible pattern: Coercion and strategic learning can be sufficient on their own, and occur in our cases in iterations. Alternatively, they can operate in combination, in which case strategic learning follows coercion. We did not find evidence of US policymakers introducing safeguard because they believed they had a moral obligation to do so. What do these findings tell us?

Coercion

Coercion does not seem to be a blunt sword when it comes to holding powerful states to account for extraterritorial human rights violations. In two of our cases, coercion was either the sole mechanism (mass surveillance) or one of two mechanisms (torture) that led to safeguards.

Material sanctions are not as powerless as one might have expected. Generally speaking, it is economically weak states that face sanctions if they do not adhere to the human rights clauses of preferential trade agreements (Hafner-Burton, 2005). Similarly, only weaker states can be threatened with a possible refusal of their bid to join an IO if they do not fulfill the IO's human

rights standards (Schimmelfennig and Sedelmeier, 2004). Nevertheless, as our case study on mass surveillance has shown, under certain conditions even powerful states like the US are vulnerable to human rights-related sanctions. One such condition is that the actor imposing sanctions has the ability to inflict significant harm on the target actor (Bapat et al., 2013): In the surveillance case, the US government depended on access to the data of foreign citizens stored by US technology companies, which meant that ignoring their demands would have been prohibitively costly.

Shaming can also work under certain conditions. Shaming certainly faces challenges when it comes to pressurizing powerful states into compliance with extraterritorial human rights obligations. For one thing, it is generally easier to mobilize the domestic public if the rights of nationals rather than those of foreigners are violated. Powerful states also generally depend less on their reputation, as they have hard power resources at their disposal. However, as we have shown in the torture case, shaming can still be effective. Specifically, it was possible to build a powerful campaign, even though it was foreigners' rights that were violated, because the violations involved physical harm that could be easily visualized and because of torture's status as a taboo – qualities that are believed to make rights violations particularly suitable for a public campaign (Barnes, 2017; Keck and Sikkink, 1998: 205).

In our cases, litigation did not prove to be an effective mechanism for holding powerful states accountable for extraterritorial human rights violations. Notwithstanding emerging case law that ascribes extraterritorial human rights obligations to states, domestic courts are frequently reluctant to issue judgments on foreign policy issues and, moreover, tend to have limited access to classified information (Setty, 2017). Furthermore, because the extraterritorial application of human rights conventions is still contested, judges have a certain margin to deny jurisdiction (Andresen, 2016). In the case of the US, it is mostly in exceptional cases where rights violations take place on territory with a complex legal status that courts are inclined to claim jurisdiction and issue judgments that demand safeguards for foreigners – as was the case with safeguards against arbitrary detention for terrorist suspects detained in Guantánamo Bay (US Supreme Court, 2006). Foreign courts can technically invoke universal jurisdiction, but they rarely do so for political reasons. Nonetheless, our analysis suggests that litigation can have indirect effects. As the surveillance case has shown, 'complicit' third parties may sometimes be vulnerable to litigation and may forward the pressure to the target state.

Strategic Learning

Strategic learning also seems to be an important mechanism for prompting powerful states to introduce safeguards to ensure extraterritorial protection of human rights. In two cases it was either the only mechanism (targeted killing) or one of two mechanisms acting together (torture) that proved effective. US policymakers processed causal arguments, weighed the anticipated future costs and benefits of safeguards, and took decisions based on those considerations (see also Abel, 2018). The establishment of safeguards cannot therefore be solely explained as a response to coercive pressure. In the cases in which we found strategic learning, we can confirm the expectation that the credibility of the messenger and the perceived risk that negative consequences will materialize in the event of inaction influence political actors' readiness to engage with strategic arguments and act upon them (Bapat et al., 2013: 89–90; Haas, 2004). In both the targeted killing and the torture case we observed that it made a difference that the predictions of the actors who put forward strategic arguments were credible. Moreover, it was important that the arguments in favor of safeguards came not only from NGOs but also from actors within the establishment who were well respected by progressives and conservatives alike.

Beyond that, there seems to be an interesting dynamic between strategic learning and coercion. In the torture case, coercion (in the form of shaming) triggered the first reform but also paved the way for strategic learning by providing an external shock that reverberated even after public attention faded. In this sense, initial coercion sensitized policymakers to the possibility that potential future coercive action might jeopardize their long-term strategic goals.

Moral persuasion

One of the most interesting findings of our analysis is that the moral persuasion mechanism did not operate as expected (see also section 2.1 of the appendix). In all cases, safeguards were introduced as the result of a rational cost–benefit analysis – either in direct response to urgent pressure or in anticipation of potential future negative consequences. This is to some extent surprising, given that being a country that values human rights is clearly part of the US identity. Moreover, all safeguards activated by the executive branch were issued by the Obama administration, which was, at least rhetorically, strongly committed to the protection of human rights.[37] So why then did we not find evidence of US policymakers introducing extraterritorial human rights safeguards because they believed that this was the right thing to do?

One explanation could be that US policymakers have only partially internalized the human rights norm and believe that it has to recede behind the security norm if the two are in conflict (Sikkink, 2013). Another explanation could be that US policymakers have internalized the idea that they have human rights obligations towards their own citizens, but are less inclined to extend them to foreigners outside of US territory. Moral persuasion may also have been crowded out by the other mechanisms. In all cases, policymakers faced strategic arguments or immediate pressure to act, in addition to moral arguments. The latter did not prove decisive, but policymakers were swayed by strategic argument or bowed to pressure – which is in line with findings that the impact of norms can be undercut by extrinsic motivation (Terechshenko et al., 2019).

That moral considerations were not the driver behind the introduction of safeguards may also help explain why in two out of three cases the safeguards enacted were not far-reaching. Those against torture are the most advanced, because they unequivocally ban all forms of it. Safeguards against indiscriminate targeted killing remained vague and were partially rolled back by the Trump administration, while many critics question the effectiveness of the safeguards against foreign mass surveillance. What is more, if the US establishes safeguards, it does not consistently use human rights language, so as not to legally or rhetorically entrap itself. Nonetheless, the safeguards are still important, both in their own right and because they are evidence of US policymakers rhetorically recognizing obligations towards foreigners abroad. Had there been a deep conviction among US policymakers that extraterritorial human rights violations are generally wrong and that it is morally appropriate to prevent their occurrence, we should have seen more far-reaching safeguards and more explicit references to human rights obligations.

This does not imply, however, that the emerging norm that states have extraterritorial human rights obligations did not play a role in the process leading up to the establishment of safeguards. What we have been able to show is that the norm did not have a constitutive effect, in the sense that norm entrepreneurs failed to convince US policymakers of the inherent value of that norm and the need to act in accordance with it. Even though individuals in the inner circle, like Senator McCain in the torture case, may have honestly believed that certain violations were inherently wrong, it was generally strategic, not moral, arguments that resonated with policymakers most. Nonetheless, the norm that states have human rights obligations also vis-à-vis foreigners beyond their borders had an important regulative effect in that it influenced US policymakers' cost–benefit calculations. The fact that other relevant actors believed in the norm

mattered because it constrained the range of justifiable policy options. Only because relevant others believed in the norm was it possible to, for instance, shame the US government for its norm-violating behavior.[38]

Last but not least, moral considerations provided motivation for actors to develop intervention strategies in the first place. In the case of torture, many actors who were instrumental in organizing a powerful shaming campaign truly believed that the ban on torture also applied to interactions with non-citizens beyond a state's borders. In the case of targeted killing, human rights groups were equally motivated by a commitment to a truly universal notion of human rights, but uttered mostly strategic arguments, as they believed that the latter resonated with policymakers. In the case of surveillance, finally, privacy activists were driven by the belief that it was wrong for the US to disregard the privacy rights of foreign citizens beyond their borders in foreign surveillance operations (see section 2.2 of the appendix).

6. Conclusion

In this article we have investigated why the US has introduced safeguards designed to prevent or at least attenuate the harm to foreigners abroad caused by their counterterrorism measures. We have searched for evidence of three mechanisms – moral persuasion, strategic learning and coercion – but found that only the latter two have had a direct effect on US policymakers. Hence, although some US policymakers had moral concerns about the harm afflicted to foreign citizens, cost–benefit calculations were key, either as a response to direct pressure or in anticipation of potential future pressure. This finding may be sobering for those who believe that the best way to make states commit to and comply with human rights norms is to change the normative convictions of policymakers. Nonetheless, our findings suggest that the norm that states should respect the human rights of foreigners beyond their borders not only provided motivation for actors to intervene with policymakers but is sufficiently strong to affect policymakers' cost–benefit analysis.

We see three avenues for future research. First, we should learn more about the conditions under which the mechanisms selected for this study operate, but also search for alternative mechanisms. Future research should therefore compare cases in which the US has introduced safeguards with cases in which it has not. One could also consider deviant cases (Seawright,

2016) and consider cases in which the US did not commit rights violations but nevertheless established safeguards to ascertain whether emulation has been at play.

Second, future research should investigate whether our findings are generalizable beyond the US. The US displays certain traits of a hard case. It is one of few established democracies that, save for limited exceptions, dismisses the idea of extraterritorial human rights obligations; moreover, it is generally powerful enough to be able to fend off pressure from other actors. Whether our findings really tell us anything about the behavior of other established democracies needs to be systematically investigated, though. One could therefore consider cases in which the mechanisms' starting and endpoints are also present, but that differ from the US regarding theoretically relevant criteria such as power, domestic institutions, or political culture - the United Kingdom's anti-torture safeguards or Israel's guidelines for targeted killing operations, for example. China's guidelines for foreign investments or Saudi Arabia's complaint mechanism for civilian victims of its operation in Yemen might also be looked into to explore to what extent the findings might apply to autocracies.

Third, we have concentrated on the rules themselves and not on their implementation. Although the emergence of safeguards is important in itself, it is, especially from a victim's perspective, also necessary to examine whether the rules are actually implemented or whether they are paper tigers intended to fend off calls for more extensive reforms. Future research should therefore look into whether and on what conditions the rules that have emerged have been implemented and how implementation can be strengthened.

7. Bibliography

Abel RL (2018) *Law's Wars: The Fate of the Rule of Law in the US 'War on Terror'*. Cambridge: Cambridge University Press.

American Civil Liberties Union (2019) About - The Torture Database. Available at: https://www.thetorturedatabase.org/about (accessed 01 July 2019).

Amnesty International (2004) USA: Pattern of Brutality and Cruelty: War Crimes at Abu Ghraib. Available at: https://www.amnesty.org.uk/press-releases/usa-pattern-brutality-and-cruelty-war-crimes-abu-ghraib (accessed 01 July 2019).

Amnesty International (2013) Log off, Mr President. On file with the authors.

Andresen J (2016) Due Process of War in the Age of Drones. *Yale Journal of International Law* 41(1): 155-188.

Argyris C and Schön DA (1978) *Organizational learning: A theory of action perspective*. Reading: Addison-Wesley.

Bapat NA, Heinrich T, Kobayashi Y and Morgan TC (2013) Determinants of sanctions effectiveness: Sensitivity analysis using new data. *International Interactions* 39(1): 79-98.

Barnes J (2017) *A Genealogy of the Torture Taboo*. London: Routledge.

Beach D and Pedersen R (2019) *Process-Tracing Methods: Foundations and Guidelines*. Ann Arbor: University of Michigan Press.

Beach D and Rohlfing I (2018) Integrating Cross-Case Analyses and Process Tracing in Set-Theoretic Research and Parameters of Debate. *Sociological Methods & Research* 47(1): 3-36.

Bennett A and Checkel JT (2015) Process tracing: From philosophical roots to best practices. In: Bennett A and Checkel JT (eds) *Process tracing: From metaphor to analytic tool*. Cambridge: Cambridge University Press, pp. 3-37.

Bignami F and Resta G (2018) Human Rights Extraterritoriality: The Right to Privacy and National Security Surveillance. In: Benvenisti E and Nolte G (eds) *Community Interests Across International Law*. Oxford: Oxford University Press, pp. 357-380.

Brown A and Newport F (2013) In U.S., 65% Support Drone Attacks on Terrorists Abroad, *Gallup*, 25 March. Available at: https://news.gallup.com/poll/161474/support-drone-attacks-terrorists-abroad.aspx (accessed 07 December 2020).

Brown I, Halperin MH, Hayes B, Scott B and Vermeulen M (2015) Towards Multilateral Standards for Surveillance Reform. Available at: https://cihr.eu/wp-content/uploads/2015/01/Brown_et_al_Towards_Multilateral_2015.pdf (accessed 08 December 2020).

Center for Civilians in Conflict (2013) Statement of shared concerns regarding US drone strikes and targeted killings: Follow up letter. Available at: https://civiliansinconflict.org/press-releases/statement-shared-concerns-regarding-us-drone-strikes-targeted-killings-follow-letter/ (accessed 08 December 2020).

Checkel JT (2006) Tracing causal mechanisms. *International Studies Review* 8(2): 362-370.

CIA Inspector General (2004) Special Review - Counterterrorism Detention and Interrogation Activities (September 2001 - October 2003). Available at: https://www.cia.gov/library/readingroom/docs/0005856717.pdf (accessed 08 December 2020).

Coleman JS (1986) Social Theory, Social Research, and a Theory of Action. *American Journal of Sociology* 91(6): 1309-1335.

Court of Justice of the EU (2015) Judgment of the Court (Grand Chamber), Case C-362/14 Maximilian Schrems v Data Protection Commissioner, October 6. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62014CJ0362&from=EN (accessed 08 December 2020).

Court of Justice of the EU (2020): Judgment of the Court (Grand Chamber), Case C-311/18 Data Protection Commissioner v Facebook Ireland Ltd. and Maximillian Schrems, July 16, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62018CJ0311 &from=EN (accessed 07 December 2020).

Davis LE, McNerney M and Greenberg MD (2016) Clarifying the rules for targeted killing: An analytical framework for policies involving long-range armed drones. Available at: https://www.rand.org/content/dam/rand/pubs/research_reports/RR1600/RR1610/RAND_RR1610.pdf (accessed 25 June 2019).

Department of the Navy (2004) Statement for the Record: Office of General Counsel Involvement in Interrogation Issues. Available at: https://www.thetorturedatabase.org/files/foia_subsite/mora_memo_july_2004_0.pdf (accessed 01 July 2019).

Donno D and Neureiter M (2018) Can human rights conditionality reduce repression? Examining the European Union's economic agreements. *The Review of International Organizations* 13(3): 335-357.

Donohue LK (2015) High Technology, Consumer Privacy, and U.S. National Security. *American University Business Law Review* 4(1): 11-48.

Downs GW, Rocke DM and Barsoom PN (1996) Is the Good News about Compliance Good News about Cooperation? *International Organization* 50(3): 379-406.

Duffy H (2018) *Strategic human rights litigation: Understanding and maximising impact*. Oxford: Bloomsbury Publishing.

European Center for Constitutional and Human Rights (2019) The Case of Bünyamin E. in Pakistan. Available at: https://www.ecchr.eu/en/case/the-case-of-buenyamin-e-in-pakistan/ (accessed 06 April 2019).

Farrell H and Newman A (2016) The transatlantic data war: Europe fights back against the NSA. *Foreign Affairs* 95(1): 124-133.

Finnemore M and Sikkink K (1998) International Norm Dynamics and Political Change. *International Organization* 52(4): 887-917.

Gammeltoft-Hansen T and Vedsted-Hansen J (2017) Introduction: human rights in an age of international cooperation, In Gammeltoft-Hansen T and Vedsted-Hansen J (eds) *Human Rights and the Dark Side of Globalisation: Transnational law enforcement and migration control*. London: Routledge, pp. 1-23.

Gibney M, Tomaševski K and Vedsted-Hansen J (1999) Transnational state responsibility for violations of human rights, *Harvard Human Rights Journal* 12: 267–295.

Goertz G (2017) *Multimethod research, causal mechanisms, and case studies: An integrated approach*. Princeton: Princeton University Press.

Goodman R and Jinks D (2013) Social Mechanisms to Promote International Human Rights: Complementary or Contradictory? In: Risse T, Ropp SC and Sikkink K (eds) *The Persistent Power of Human Rights: From Commitment to Compliance*. Cambridge: Cambridge University Press, pp.103-122.

Greenfield D and Hausheer SA (2014) Do Drone Strikes in Yemen Undermine US Security Objectives? Available at: https://www.files.ethz.ch/isn/184589/Drone_Strikes_in_Yemen.pdf (accessed: 08 December 2020).

Greenwald G (2014) *No place to hide: Edward Snowden, the NSA, and the US surveillance state.* New York: Metropolitan Books.

Greenwald G and MacAskill E (2013) NSA Prism program taps in to user data of Apple, Google and others. *The Guardian*, 7 June. Available at: https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data (accessed: 08 December 2020).

Grobe C (2010) The power of words: Argumentative persuasion in international negotiations. *European Journal of International Relations* 16(1): 5-29.

Haas P (2004) When does power listen to truth? A constructivist approach to the policy process. *Journal of European Public Policy* 11(4): 569–92.

Hafner-Burton EM (2005) Trading human rights: How preferential trade agreements influence government repression. *International Organization* 59(3): 593-629.

Hall PA (2013) Tracing the Progress of Process Tracing. *European Political Science* 12(1): 20-30.

Hanley CJ (2003) AP Enterprise: Former Iraqi detainees tell of riots, punishment in the sun, good Americans and pitiless ones. *Associated Press*, 1 November. Available at: https://web.archive.org/web/20140222024344/http://legacy.utsandiego.com/news/world/iraq/20031101-0936-iraq-thecamps.html (accessed 07 December 2020).

Hersh SH (2004) Torture at Abu Ghraib. American soldiers brutalized Iraqis. How far up does the responsibility go? 15 May. Available at: https://www.newyorker.com/magazine/2004/05/10/torture-at-abu-ghraib (accessed 08 December 2020).

Heupel M (2020) Indirect Accountability for Extraterritorial Human Rights Violations. *International Studies Perspectives* 21(2): 172-97.

Heupel M and Hirschmann G (2017) Conceptual Framework. In: Heupel M and Zürn M (eds) *Protecting the Individual from International Authority. Human Rights in International Organizations*, Cambridge: Cambridge University Press, 40-65.

Human Rights First (2008) Top Interrogators Declare Torture Ineffective in Intelligence Gathering. Available at: https://www.humanrightsfirst.org/2008/06/24/top-interrogators-declare-torture-ineffective-in-intelligence-gathering (accessed 08 December 2020).

Human Rights Watch (2013) Between a Drone and Al-Qaeda. Available at: https://www.hrw.org/report/2013/10/22/between-drone-and-al-qaeda/civilian-cost-us-targeted-killings-yemen (accessed 27 April 2021).

Keck ME and Sikkink K (1998) *Activists beyond borders: Advocacy networks in international politics*. Ithaca: Cornell University Press.

King H (2009) The Extraterritorial Human Rights Obligations of States. *Human Rights Law Review* 9(4): 521-556.

Klotz A (1995) Norms reconstituting interests: global racial equality and US sanctions against South Africa. *International Organization* 49(3): 451-478.

Krain M (2012) J'accuse! Does naming and shaming perpetrators reduce the severity of genocides or politicides? *International Studies Quarterly* 56(3): 574-589.

Lawrence C (2013) Drone Attack Survivors Take Their Story to Washington. Available at: http://security.blogs.cnn.com/2013/10/30/drone-attack-survivors-take-their-story-to-washington/ (accessed 06 March 2019).

Levy S (2014) How the NSA Almost Killed the Internet. *The Wired*, 1 July. Available at: https://www.wired.com/2014/01/how-the-us-almost-killed-the-internet/ (accessed 15 February 2018).

Mayer J (2006) The memo. How an internal effort to ban the abuse and torture of detainees was thwarted. 19 February. Available at: https://www.newyorker.com/magazine/2006/02/27/the-memo (accessed 30 April 2021).

Murdie AM and Davis DR (2012) Shaming and blaming: Using events data to assess the impact of human rights INGOs. *International Studies Quarterly* 56(1): 1-16.

New America (2019) Drone Strikes: Pakistan. Available at: https://www.newamerica.org/in-depth/americas-counterterrorism-wars/us-targeted-killing-program-yemen/ (accessed 06 March 2019).

Office of Legal Counsel (2002a) Memorandum for Alberto R. Gonzales - Re: Standard of Conduct for Interrogation under 18 U.S.C. §§ 2340-2340A. Available at: https://nsarchive2.gwu.edu/NSAEBB/NSAEBB127/02.08.01.pdf (accessed 30 June 2019).

Office of Legal Counsel (2002b) Memorandum for William J. Haynes, II - Re: Potential Legal Constraint Applicable to Interrogations of Persons Captured by U.S. Armed Forces in Afghanistan. Available at: https://www.justice.gov/sites/default/files/olc/legacy/2009/08/24/memo-constraints-feb2002.pdf (accessed 30 June 2019).

Rascoff SJ (2016) Presidential Intelligence. *Harvard Law Review* 129(3): 633-717.

Reinold T and Zürn M (2014) "Rules about Rules" and the Endogenous Dynamics of International Law: Dissonance Reduction as a Mechanism of Secondary Rule-making. *Global Constitutionalism* 3(2): 236-273.

Risse T (2000) "Let's argue!" Communicative action in world politics. *International Organization* 54(1): 1-39.

Risse T, Ropp SC and Sikkink K (1999) *The power of human rights: International norms and domestic change*, Cambridge: Cambridge University Press.

Roth K (2013) the NSA's Global Threat to Free Speech. 18 November. Available at: https://www.hrw.org/news/2013/11/18/nsas-global-threat-free-speech (accessed 15 April 2021).

Seawright J (2016) The Case for Selecting Cases that are Extreme on the Independent Variable. *Sociological Methods and Research* 45(3): 493-525.

Schimmelfennig F (2015) Efficient process tracing. Analyzing the causal mechanisms of European integration. In: Bennett A and Checkel JT (eds) *Process tracing: From metaphor to analytic tool*. Cambridge: Cambridge University Press, pp. 98-125.

Schimmelfennig F and Sedelmeier U (2004) Governance by conditionality: EU rule transfer to the candidate countries of Central and Eastern Europe. *Journal of European public policy* 11(4): 661-679.

Schneier B (2015) *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*. New York: W. W. Norton & Company.

Senate Select Committee on Intelligence (2014) Committee Study of the Central Intelligence Agency's Detention and Interrogation Program. Senate Report 113-288, 9 December. Available at: https://fas.org/irp/congress/2014_rpt/ssci-rdi.pdf (accessed 08 December 2020).

Setty S (2015) The United States. In: Roach K (eds) *Comparative Counter-Terrorism Law*. Cambridge: Cambridge University Press, pp. 49-77.

Setty S (2017) *National security secrecy: Comparative effects on democracy and the rule of law*. Cambridge: Cambridge University Press.

Sikkink K (2013) The United States and torture: Does the spiral model work? In: Risse T, Ropp SC and Sikkink K (eds) *The Persistent power of human rights: from commitment to compliance*. Cambridge: Cambridge University Press, pp.145-163.

Simmons BA (2009) *Mobilizing for human rights: International law in domestic politics*. Cambridge: Cambridge University Press.

Skogly S (2006) *Beyond national borders: States' human rights obligations in international cooperation*. Oxford: Intersentia.

Terechshenko Z, Crabtree C, Eck K and Fariss CJ (2019) Evaluating the influence of international norms and shaming on state respect for rights: an audit experiment with foreign embassies. *International Interactions*, 45(4), 720-735.

The Bureau of Investigative Journalism (2019) Drone Warfare. Available at: https://www.thebureauinvestigates.com/projects/drone-war (accessed 25 June 2019).

The White House (1981) Executive Order 12333 – United States intelligence activities. *Federal Register*, Vol. 46, No. 235, 8 December. Available at: https://www.archives.gov/federal-register/codification/executive-order/12333.html (accessed 08 December 2020).

The White House (2005) President's Statement on Signing of H.R. 2863, the "Department of Defense, Emergency Supplemental Appropriations to Address Hurricanes in the Gulf of Mexico, and Pandemic Influenza Act, 2006". Available at: https://georgewbush-whitehouse.archives.gov/news/releases/2005/12/20051230-8.html (accessed 08 December 2020).

The White House (2009) Executive Order 13491 – Ensuring Lawful Interrogations. *Federal Register*, Vol. 74, No. 16, 27 January. Available at: https://obamawhitehouse.archives.gov/the-press-office/ensuring-lawful-interrogations (accessed 16 November 2020).

The White House (2013a) Procedure for Approving Direct Action Against Terrorist Targets Located Outside the United States and Areas of Active Hostilities, Presidential Policy Directive. Available at: https://fas.org/irp/offdocs/ppd/ppg-procedures.pdf (accessed 06 April 2019).

The White House (2013b) Remarks by the President at the National Defense University. Available at: https://obamawhitehouse.archives.gov/the-press-office/2013/05/23/remarks-president-national-defense-university (accessed 06 March 2019).

The White House (2014a) Presidential Policy Directive/PPD-28 - Signals Intelligence Activities, 17 January. Available at: https://fas.org/irp/offdocs/ppd/index.html (accessed 08 December 2020).

The White House (2014b) Press Conference by the President. Available at: https://obamawhitehouse.archives.gov/the-press-office/2014/08/01/press-conference-president (accessed 28 June 2019).

The White House (2016) Executive Order 13732 – United States Policy on Pre- and Post-Strike Measures To Address Civilian Casualties in U.S. Operations Involving the Use of Force. *Federal Register*, Vol. 81, Issue 130, 7 July. Available at: https://www.federalregister.gov/documents/2016/07/07/2016-16295/united-states-policy-on-pre-and-post-strike-measures-to-address-civilian-casualties-in-us (accessed 06 April 2019).

The White House (2019) Executive Order 13862 - Revocation of Reporting Requirement. Federal Register, Vol. 84, No. 47, 11 March. Available at: https://www.whitehouse.gov/presidential-actions/executive-order-revocation-reporting-requirement/ (accessed 04 December 2020).

Timberg C (2013) Major tech companies unite to call for new limits on surveillance. *The Washington Post*, 9 December. Available at: https://www.washingtonpost.com/business/technology/major-tech-companies-unite-to-call-for-new-limits-on-surveillance/2013/12/08/ 530f0fd4-6051-11e3-bf45-61f69f54fc5f_story.html (accessed 29 April 2021).

UN Human Rights Committee (2014) Concluding Observations on the Fourth Periodic Report of the United States of America. Available at: http://www.refworld.org/docid/5374afcd4.html (accessed 15 February 2018).

UN Human Rights Council (2010): Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Philip Alston, A/HRC/14/24/Add.6. Available at: https://www2.ohchr.org/english/bodies/hrcouncil/docs/14session/A.HRC.14.24.Add6.pdf (accessed 03 May 2021).

US Congress (2001) Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA Patriot Act) of 2001. Public Law 107-56, 26 October. Available at: https://www.congress.gov/107/plaws/publ56/PLAW-107publ56.pdf (accessed 15 March 2017).

US Congress (2005) Detainee Treatment Act of 2005, Title X of the Department of Defense, Emergency Supplemental Appropriations to Address Hurricanes in the Gulf of Mexico, and Pandemic Influenza Act, 2006. Public Law 109–148, 30 December. Available at: https://www.govinfo.gov/content/pkg/COMPS-489/pdf/COMPS-489.pdf (accessed 08 December 2020).

US Congress (2008) Foreign Intelligence Surveillance Act of 1978 Amendments Act of 2008. Public Law 110-261, 10 July. Available at: https://www.congress.gov/bill/110th-congress/house-bill/6304 (accessed 15 February 2018).

US Congress (2015) National Defense Authorization Act for fiscal year 2016. Public Law 114-92, 25 November. Available at: https://www.congress.gov/114/plaws/publ92/PLAW-114publ92.pdf (accessed 08 December 2020).

US Department of Commerce (2016) EU-U.S. Privacy Shield Framework Principles Issued by the U.S. Department of Commerce. Available at: https://www.privacyshield.gov/servlet/servlet.FileDownload?file=015t00000004qAg (accessed 28 November 2016).

US Supreme Court (2006) Opinion, No. 05–184 Hamdan v Rumsfeld, Secretary of Defense, et al., June 29. Available at: https://www.supremecourt.gov/opinions/05pdf/05-184.pdf (accessed 09 December 2020).

Vandenhole W and Van Genugten W (2015) Introduction: An Emerging Multi-Duty-Bearer Human Rights Regime? In: Vandenhole W (ed), *Challenging Territoriality in Human Rights Law: Building Blocks for a Plural and Diverse Duty-Bearer Regime*. London and New York: Routledge, pp. 1-11.

8. Notes

[1] Strategic learning thus comes close to Argyris and Schön's (1978) single-loop learning, because policymakers do not alter their normative convictions but merely adjust their behavior to increase policy effectiveness.
[2] On equifinality generally see Beach and Pedersen (2019: 6).
[3] Interview with Elizabeth Grimm Arsenault, Georgetown University, Washington D.C., March 19, 2019.
[4] Interview with Elizabeth Grimm Arsenault.
[5] Interview with Benjamin Wittes, Senior Fellow, Brookings Institution, Washington D.C. March 19, 2019.
[6] Interview with Philip Zelikow, Executive Director of the National Commission on Terrorist Attacks Upon the United States (9/11 Commission), Washington D.C., April 8, 2019.
[7] Interview with Mark Fallon, former Chair of the US Government's High Value Detainee Interrogation Group Research Committee, via Skype, May 3, 2019.

[8] Interview with Benjamin Wittes.

[9] Interview with Elisa Massimino, former President and CEO of Human Rights First, Washington D.C., March 28, 2019.

[10] Interview with John Bellinger, Senior Associate Counsel to President Bush and Legal Advisor to the National Security Council, Washington D.C., April 9, 2019.

[11] Interview with Elisa Massimino.

[12] Interview with Michael Posner, former Assistant Secretary of State for the Bureau of Democracy, Human Rights and Labor, via Skype, July 1, 2019.

[13] Interview with Mark Fallon.

[14] Interview with Elisa Massimino.

[15] Interview with Colonel Steven Kleinman, Military Intelligence Officer, Washington D.C., April 2, 2019.

[16] 'Lethal action' is the technical term for targeted killing, often featured in Obama administration documents.

[17] Interview with Patricia Stottlemyer, Impact Litigation Attorney at Human Rights First, Washington D.C., March 26, 2019.

[18] Interview with Kenneth Anderson, American University, Washington D.C., March 21, 2019.

[19] Interview with Larry Lewis, former Senior Advisor to the Department of State's Assistant Secretary for Democracy, Human Rights, and Labor, Washington D.C., April 3, 2019.

[20] Interview with Luke Hartig, former Senior Director for Counterterrorism at the National Security Council, via Skype. July 2, 2019.

[21] Interview with Larry Lewis.

[22] Interview with Kenneth Anderson.

[23] Interview with Sarah Holewinski, Executive Director, Center for Civilians in Conflict, Washington D.C., March 22, 2019.

[24] Interview with Larry Lewis.

[25] Interview with Luke Hartig.

[26] Interview with Larry Lewis.

[27] Interview with Kenneth Anderson.

[28] Interview with Larry Lewis.

[29] This section draws on Heupel (2020).

[30] Interview with Amie Stepanovich, US Policy Manager, Access Now, Washington D.C., February 28, 2017.

[31] Interview with Greg Nojeim, Senior Counsel, Center for Democracy & Technology, Washington D.C., February 28, 2017.

[32] Interview with Robert Litt, former General Counsel to the Office of the Director of National Intelligence, Washington D.C., March 14, 2019.

[33] Interview with a US Senate staffer, Washington D.C., February 27, 2017.

[34] Interview with a senior EU official, Florence, May 5, 2017.

[35] Interview with Neema Guliani, Senior Legislative Counsel, American Civil Liberties Union, Washington D.C., March 13, 2019.

[36] Following the Snowden revelations, President Obama appointed a Review Group on Intelligence and Communications Technologies. It is unlikely, however, that the recommendations provided by this group would have been followed had there not been pressure from the tech industry.

[37] The two laws that we cover in the torture case study were enacted by Congresses with Republican majorities.

[38] On the distinction between the constitutive and regulative effect of norms see Klotz (1995).

<div align="center">

Article 2:

When Identity Meets Strategy

– The development of British and German anti-torture policies since 9/11

</div>

Abstract

Since 9/11, a lot of research has been done on US interrogation and detention practices in Guantanamo, Abu Ghraib, and other black sites around the world, but we still know little about the respective involvement of other traditionally liberal states' intelligence agencies and their evolving perspectives on torture-related policies for foreigners abroad. Particularly the UK and Germany pose interesting cases in this regard, as in 2010, despite similar levels of public and political pressure following revelations about their indirect involvement in CIA operations, the two states took different strategic decisions when it came to implementing or forgoing new extraterritorial human rights safeguards. While the UK introduced a new intelligence guidance for interrogations overseas, the German government opted for policy-continuance, which raises the question why the two states embarked on different policy trajectories, even if they found themselves in contextually similar situations and were subjected to the comparable accountability measures. By drawing on Rationalism and taking inspiration from normative literature, the article addresses this conundrum by not only clearly outlining the states' differing strategic preferences, but also by dissecting the multi-layered composition of these interests. As a result, the paper delineates how strategic constraints pertaining to the states' national, international, or political elite level affect the decision-makers' cost-benefit analyses and thus their respective policy responses.

Keywords: Human Rights, Counterterrorism, Foreign Policy, Convention against Torture, Identity, Rational-Choice, UK, Germany

Introduction

Why should states refrain from torturing foreigners abroad? A question seemingly easy to answer and yet one that often sparks debates, as illustrated by the different outcomes of British and German policy-making during the first decade of the War on Terror. Though both countries similarly pride themselves with being democratic, Western advocates for human rights, both states, in the aftermath of 9/11, grew complicit to the CIA Detention and Interrogation Program

by among others enabling rendition flights via their respective air spaces and by providing vital information and question sets to the US agency (European Parliament 2006). Although individual numbers and official accounts of breaches vary, the public and international backlash against the states' intelligence services manifested in similar ways: Both countries were identified in the Dick Marty Report by the EU Parliament for their active facilitation of logistics as well as in the exchange of information despite their awareness of the CIA's detainee mistreatment (European Parliament 2006); the German Bundesnachrichtendienst (BND) and the British Secret Intelligence Service (SIS) were publicly shamed by renown NGOs for their norm-violating behaviors (Amnesty International 2009; Human Rights Watch 2006), which were confirmed as indirect violations of the Convention against Torture by inquiries from their respective parliaments; and both countries experienced moderate waves of public protest throughout the early 2000s.

Notwithstanding the similar degrees of public and international backlash against their intelligence agencies' complicity in the CIA program, the policy reactions of each country differed greatly: While in 2010 David Cameron issued a new interrogation guidance for overseas missions, the German government remained unmoved, successfully withstanding any demands for reforming German interrogation operations. Given the particular sensitivity of torture allegations and the two states' carefully crafted self-image of being liberal, Western human rights advocates, one might think that both states would have reacted comparably to the relatively same pressure; or at minimum that the Merkel Administration would have been more inclined to embrace accountability in order to avoid any comparisons to the country's dark past during WWII. This is not to say that Cameron's guidance constituted a benevolent panacea against torture, in fact it has been severely criticized as window-dressing and has been since updated (see "The Principles" HM Government 2019); nonetheless it suggests that the British government perceived a policy-change as strategically beneficial, while his German counterpart seemingly did not. Hence, in light of this apparent contradiction, it is necessary to explain why the two states embarked on different policy trajectories in order to better understand the effectiveness of shaming and other traditional means of accountability.

When looking at Human Rights literature, scholars explain differences in state behavior by scrutinizing the level of external influence via coercion or persuasion efforts (Hawkins 2004), the actors' degree of norm internalization (Sikkink 2013), and the role of international organizations in holding states accountable (Creamer and Simmons 2018; Johansen 2020). In the field of Intelligence Studies, in turn, scholars attribute the states' diverse stances on torture

to the country-specific evolution of counterterrorism paradigms in the post-9/11 era (Luban 2007), as well as to the general discussion about the effectiveness of torture itself (Brimbal et al. 2019). In the same context, scholars focused on the development of democratic states' torture policies since 9/11, less though on the creation of extraterritorial protection safeguards (see e.g. Liese 2009; Hafner-Burton and Shapiro 2010; Blakely and Raphael 2017). Though all three strands of literature offer valuable insight into states' motivation for (non-) compliance with the international ban on torture, neither of them can single out an observable cause for the countries' diverging conclusions as the two states respond for instance to the same international organizations, share similar counterterrorism paradigms, and have also faced similar degrees of (inter-)national pressure.

In light of this research gap, the manner in which this paper proceeds is twofold: First, it briefly analyzes the British and German decision-making processes in order to gain an overview of *how* the countries' strategic arguments and cost-benefit analyses diverge. Based on these findings, the study then takes inspiration from normative literature and insights from research on state identity to delve into the multi-layered composition of strategic interests and constraints exploring *why* the decision-makers' assessments differ. Specifically, this paper scrutinizes how different levels of the British and German political landscape have impacted the states' strategic positioning and preferences towards torture in the post-9/11 era, with a special focus being put on the countries' intelligence services. In doing so, the paper identifies three types of constraints and traces them through the strategic rationales displayed in diverse primary sources as well as 21 semi-structured interviews that have been conducted with relevant British and German experts, politicians, and stakeholders.

The analysis' outcome contributes to the academic debate in three ways. To begin with, it (1) enhances the empirical knowledge about states' strategic policy-making processes in the aftermath of security-related human rights violations. By exploring the British and German cases from a rationalist angle enriched by insights from normative literature, however, the (2) article emphasizes the interplay of a state's identity and strategic preferences, accounting thus for country-specific multi-layered strategic constraints. Yet, it differs from other studies with a similar approach by not considering identity as one-size-fits-all approach, but rather (3) scrutinizes three different dimensions of the concept as well as their respective impact on the state actors' strategic calculations. Finally, the findings are also of practical relevance, as they offer an insight in the states' interests and motivations, which is of particular importance when trying to hold accountable countries that have aided and abetted the CIA's use of torture.

Subsequently, the article follows a two-step structure. The first part traces briefly the British and German political debates between 2001 and 2010, in order to outline the cornerstones of the states' strategic (dis)interest in anti-torture safeguards for foreigners abroad. The respective results show that the UK perceived policy-change as a beneficial tool to restore its international reputation and to prevent future domestic discontent, which resulted in the creation of a a new guidance without promoting a panacea against torture. Germany, in contrast, maintained a narrative of denial in order to minimize the diverse political costs of policy-change being misinterpreted as admission of guilt; preferring instead a strategy of policy-continuance to secure its foreign-policy interests. Building on these findings and inspired by literature on state identity, the second part introduces three different dimensions of strategic constraints (*national, international, and political elite*) to subsequently explore in detail their influence on the states' strategic interests and situational assessments.

Part I: The States' Strategic Positioning

In the aftermath of 9/11, the fear of future terrorist attacks dominated the national security paradigms in most Western states, creating thus a platform for coordinated intelligence operations, which both directly and indirectly violated the prohibition of torture. While, the CIA's Enhanced Interrogation Techniques stood in stark contrast to the most basic principles of the Convention against Torture (CAT), the complicit involvement of the UK and Germany did not necessarily accumulate to a systematic form of direct mistreatment, but still enabled, assisted, and encouraged US practices (European Parliament 2006). According to official investigations, both countries turned a blind eye on the US *modus operandi*, supplying the CIA with comprehensive background information and location details of suspects, providing question sets for the ensuing interrogations, and assisting the administration with US rendition flights on multiple occasions (European Parliament 2006, 5-6, 15; Deutscher Bundestag 2009a, 434, 461; Intelligence and Security Committee 2018, 2-4, 51, 57, 83).

Over the following years, however, the political landscape and the terrorist threat assessment started to change, triggering questions regarding future collaboration between the UK, Germany, and the CIA. According to Rational Choice Theory, such changes generally require states to reassess the benefits of continued collaboration in light of new operational contexts (Snidal 2012; Hafner-Burton et al. 2017). Such considerations are especially sensitive to the volatility of public opinion, which was particularly relevant as a major scandal involving US detention and interrogation in Abu Ghraib and in Guantanamo Bay began to unfold.

Specifically, elected decision makers in the UK and Germany had to consider whether the evolving nature of the terrorist threat, as well as the weakening of key terrorist organizations still justified the high costs of being associated with increasingly scrutinized CIA practices. The following sections will briefly explore the situational background and the calculations made by policymakers in both countries as they debated whether to keep or to revise the rules governing intelligence cooperation and interrogation procedures.

United Kingdom

Upon joining the 'Coalition of the Willing' and as member to both the NATO and the 'Five Eyes'[5], the UK swiftly supported the US 'War on Terror' by sending its military and Secret Intelligence Service (SIS) to Afghanistan and Iraq; while first being publicly supported, however, the close cooperation eventually turned into a complex political balancing act. The first big scandal erupted in 2004 with the leak of very graphic photos depicting the severe cases of detainee abuse in Abu Ghraib and Guantánamo; though most perpetrators could be identified as American, the British media, NGOs, and elected officials demanded clarifications regarding the British military's role in the matter (Intelligence and Security Committee 2005, 2). The ensuing parliamentary investigation claimed not to have found any proof for British personnel violating the Geneva Conventions (Intelligence and Security Committee 2005, 29); concluding, however, that "[t]he SIS and Security Service personnel deployed to Afghanistan and Guantanamo Bay were not sufficiently trained in the Geneva Conventions, nor were they aware which interrogation techniques the UK had specifically banned in 1972" (Intelligence and Security Committee 2005, 30). Relatedly, the report also provided evidence for British intelligence personnel being witness to several cases of American CAT violations (p. 22-26) which neither interrupted nor stopped the cooperation between the two states; in total British intelligence personnel conducted and witnessed more than 2000 interviews in Afghanistan, Guantánamo, and Iraq (p.29). In parallel, the European Union criticized the UK for its complicity in the CIA rendition operations, while also accusing the SIS of sharing vital information and question sets with their American counterparts (European Parliament 2006). Despite these repeated accusations of norm-violating behavior, the public remained relatively silent forgoing any significant public pressure on the government.

---

[5] Intelligence Cooperation between Australia, Canada, New Zealand, the UK, and the US.

Having thus inherited a latently complex situation, David Cameron, upon entering office in 2010, promptly looked into various measures to address the reoccurring torture debate and to strategically avoid any further negative repercussions. In the domestic context, a new policy offered the government the opportunity to pro-actively craft a narrative, which, regardless of the actual intentions, depicted Cameron as a norm-compliant leader who was 'doing something' to investigate the allegations, and to thus protect the British "heroes" in the MI6 from the international wave of criticism and litigation (House of Commons 2010, Column 175). On the international level, in turn, the policy promised to limit further reputational harm and to reinforce the UK's positioning as a reliable, transparent, and law-abiding intelligence partner (House of Commons 2010, Column 176). Due to the similar timing of Obama's 2009 Executive Order 13491 establishing a general torture ban for US agencies, a similar UK policy was not expected to carry high costs regarding the UK-US relationship, lowering therefore the anticipation of operational or security-related costs (Interview 3). In the end, the perceived benefits of introducing the reforms outweighed the anticipated costs, paving thus the way for the so-called **"Consolidated Guidance"** [6] as well as various investigations examining the allegations against the UK. The new policy stressed the strict prohibition of torture and cruel/inhumane/degrading treatment, extended the guidelines for how to address concerns about detainee mistreatment, and increased the accountability level by making the guideline accessible to the public. Despite these improvements, however, it also gave great discretionary power to Ministers while also, among others, failing to extend the Guidance to various external cooperation partners (HM Government, 2010).

Germany

Contrary to the UK, Germany did not join the 'War on Terror' with the "Coalition of the Willing", but rather as a NATO member state; regardless, Berlin quickly saw itself confronted with strong international criticism concerning its involvement in CIA operations. The European Parliament, for instance, released multiple reports emphasizing, among others, Germany's role in various cases of rendition, abduction, and detainee mistreatment, outlining in detail how and where Germany colluded with the US intelligence community (European Parliament 2006; 2007, 15-17). On a similar note, the UN Special Rapporteur for Torture demanded Germany in 2006 to ensure that neither German authorities nor German territory be involved to illegally

---

[6] Full name: "Consolidated Guidance to Intelligence Officers and Service Personnel on the Detention and Interviewing of Detainees Overseas, and on the Passing and Receipt of Intelligence Relating to Detainees".

transfer detained individuals (European Parliament 2007, 17). Domestically, the media and NGOs called out senior politicians for publicly supporting the use of evidence obtained by torture, while also repeatedly covering the severe mistreatment of, among others, Murat Kurnaz, Mohammed Haydar Zammar, and Khaled al-Masri, whose right to be free from torture was allegedly violated by the complicit behavior of the Bundesnachrichtendienst (BND) (Deutsche Welle 2005; Frankfurter Allgemeine Zeitung 2007; Amnesty International, 2009). Though the resulting public uproar failed to translate into palpable political pressure, the opposition parties filed for a parliamentary investigation to scrutinize the German government's knowledge of and involvement in US' rendition and interrogation procedures (Deutscher Bundestag 2009a, 1-2). As a result, a parliamentary report was published in 2009, concluding that information shared by the BND did indeed lead to unlawful CIA detentions (Deutscher Bundestag 2009a, 432-442), that the German government had intentionally ignored allegations of an American base in Mannheim being used as illegal CIA detention facility (p.461), and that according to third party investigations by NGOs, German authorities opened the country's airspace for related rendition flights in over 400 cases (p. 455). However, the report also states that ultimately neither the German nor the American governments had been willing to provide any related information or evidence regarding these claims, so that the exact number could not be confirmed (p. 59).

In the meantime, the German government remained rather inactive, forgoing thus any substantive change to the rules governing German intelligence sharing and interrogation practices. Given the moderate public pressure, the anticipated electoral and political costs of policy-continuance remained relatively low, reducing thus, on the domestic level, the strategic benefits of a new policy. Conversely, the active pursuit of policy-change would likely go hand-in-hand with at least a partial recognition of CAT-violating behavior by German troops or agents, threatening thus to 'wake sleeping dogs' (Interview 18). In a similar fashion, although the mostly rhetorical reaction on the international stage exposed Germany's non-compliance with international human rights standards, similar complicity by other European member states drastically diminished the likelihood of costly material sanctions (Nowak 2010, 4-5). Instead, further concerns centered around the fact that a policy-reform could jeopardize the functioning of the close partnership between the CIA and the BND, which could not only lead to conflict with US agencies, but also counteract Germany's national security interests in the 'War on Terror' (Hannig in Deutscher Bundestag 2009a, 403). As a result, Merkel and her government did not consider policy-change to be a feasible option. Although the parliament did indeed pass a new bill increasing its oversight powers over the BND, the law did not refer to torture or

interrogation, nor did the corresponding public debates center much on detainee mistreatment (Bundesgesetzblatt 2009).

In sum, British and German complicity to the CIA Detention and Interrogation Program experienced pressure from international actors like the EU or the UN and triggered official investigations with similar outcomes, resulting in similarly strong levels of criticism from the media and NGOs. Upon deciding whether to change the rules for intelligence sharing and interrogation procedures, however, each state interpreted the situation differently, reaching ultimately opposite conclusions regarding the costs and benefits of policy-change. So why do states with such a similar situational background come to such different assessments? Why does one state perceive policy-change as a threat to their domestic and international reputation, while the other one perceives it as reputational amends? And why does one country, at a time where the US itself had already introduced new interrogation guidelines, fear upsetting their American partner, while the other one considers such a scenario as rather unlikely?

Part II: Identity Meets Strategy – The evolution of strategic constraints

The following section builds on the above findings, delving into *why* the British and German strategic assessments of the situation differ so greatly. In doing so, the first part of the section will outline basic theoretical assumptions on Rationalism and use insights from the literature on state identity to further elaborate on the concept of strategic constraints. The second part puts these theoretical postulations into test by applying them to the British and German decision-making processes.

Theory

As indicated by the name, rationalist theory assumes political actors to behave rational, meaning that based on their individual preferences, interests, and information available they assess policy choices by weighing expected benefits against potential costs in order to anticipate and to pursue the most gain-maximizing strategy to reach their goals (Fearon 1995; Simon 1997). Accordingly, rational actors are characterized by risk-adverse, evaluative, and self-interested behavior, which ultimately serves the purpose of securing the greatest benefits for themselves (Snidal 2012, 87; Epstein 2013, 293). While these notions can be applied to all actors in the political arena, the theory defines state actors as any incumbent state representative with

sufficient authority and power to speak and act for their respective country on the international stage (Alden 2017).

According to this approach, state actors proceedings are not only guided by their strategic preferences, but also by external limitations, and the principal of bounded rationality (Simon 1997; Frieden 1999, 39; Shannon et al. 2019). Correspondingly, based on national and individual interests, each actor develops intrinsic strategic preferences, which do not only reflect the perceived ideal result of each decision-making process, but also provide a ranking of less favorable options, which are decreasingly compatible with the actor's interest (Scharpf 2018, 79). In light of changing strategic environments, other actors' involvement, uneven power distributions, distinct expectations, and limited information available, however, external limitations and constraints are imposed on the actors' decision-making; forcing them eventually to not only consider their own preferences, but also the likelihood of successfully realizing them (Hafner-Burton et al. 2017, 19).

Suitably, actors attempt to maximize the utility of their choices by aligning them with their preferences regarding the outcome and the menu of alternatives, while their beliefs concerning the state of the world and the other actors' preferences inform the decision (Hafner-Burton et al. 2017, 6-7). In this context, the rational-choice model accounts for heterogeneous risk, action, and outcome preferences, that are unique to each actor and can, but ultimately do not have to be materialistically motivated (Fearson and Wendt 2002, 59). Given the model's positivist approach, however, Rationalism falls short in comprehensively explaining why exactly the actors' preferences differ or might even change over time. It is in this context that this article looks into literature on state identity to filter out potential analytical patterns that can be translated into the rationalist understanding of constraints and thus provide further insight in the development of political actors' intrinsic hierarchy of preferences. In doing so, the article takes inspiration from typically Constructivist principles, without aiming to fully marry the two schools of thought like for instance Jupille et al. (2003) or Adler and Pouliot (2011) did. Instead, the article uses insights from Hopf's (2002) notion of each (state-)actor having multiple identities to create a multi-layered model of strategic constraints, and to allow for a rationalist analysis that produces a more nuanced understanding for why the actors' preferences differ so greatly.

Following Hopf's rationale, the formation and power of state identities play a particularly important role, as they often serve as first point of orientation in an otherwise complex constellation of interests within the international system (Hopf 2002). Correspondingly, he

emphasizes that it is key to "explor[e] not only how [the] state's identities are produced in interactions with other states, but also how its identities are being produced in interaction with its own society and the many identities and discourses that constitute that society" (Hopf 2002, 294). Based on these assumptions and built on expanding insights from corresponding literature, this article deductively introduces and investigates three different types of strategic constraints, which individually account for different nuances of state identity and their respective impact on actors' preferences and options.

The first category relates to constraints emerging from the states' *national identity* characterized by a shared history, common traditions, norms and values, as well as the country's distinctive political culture. Accordingly, past events shape the collective national memory generating myths, rituals, common historic narratives, and a sense of group cohesiveness (Gillis 1994; Weedon 2012). In doing so, these experiences mark the limits and boundaries of what the people perceive as (un-)desirable actions at home and abroad, as some incidents of the past may constitute a source of national pride, while others, like international conflicts, scandals, or war atrocities can conjure strong negative recollections and a mindset of "never again" (Ryan 2012). Closely intertwined with a state's history is a country's political culture, "the traditions of a society, the spirit of its public institutions, the passions and the collective reasoning of its citizenry, and the style and operating codes of its leaders […]" (Pye 2015, 7). In other words, cultural sentiments, ideals, and attitudes spawn a web of rules, which governs the state's political system and the politicians' behavior. Classical examples for such cultural patterns would be a state's view on democratic governance, the degree of (de-) militarization, or the country's attitude towards the rule of law, yet, also components like loyalty, geographical attachments, or ethical values form an important part of a nation's expectations regarding the public life and ideals for their political realities (Duffield 1998; Pye 2015, 8). In sum, a country's national identity is heavily influenced by its history and culture, providing thus the leading political elite with a frame of suitable political actions, which in turn shape and define state preferences, while also imposing clear limitations regarding domestically unwelcomed decision-making outcomes.

The second variant of identity-based constraints scrutinized in this paper, refers to a state's *international identity* comprising of components like the country's self-perception on the global stage, its reputation among other international actors, respective dependencies and power-relations as well as the state's legitimacy claims for its international actions. Hence, the state's self-perception encompasses indicators that reveal how a state subjectively

understands, distinguishes, and envisions its own role in the international order (Manners and Whitman 2003, 383). Its international reputation, however, is determined by other states whose interpretation and assessment of the state's actions ultimately influence future cooperative opportunities (Brewster 2009, 231). As a result of these constructed and at times idealistic perceptions, alliances, rivalries, and other power structures emerge, establishing eventually a global net of dependencies and partnerships (Baldwin 2012). Moving within this net, basic rules of state interaction are marked by internationally held ideas and beliefs, whose transgression could not only trigger conflict, but also undermine the state's self-ascribed role in the international system. Hence, states often seek to legitimize their actions by referring to international structures or norms that on one hand appease to parts of the international community and on the other hand reinforce the state's self-perception (Hurd 1999). In doing so, constraints linked to a state's international identity affects a state's preferences as it factors in their actions' prohibitive costs and consequences on their international power and standing.

Finally, the third variant of constraints is linked to the identity of the state's *political elite*, meaning to government constellations, dominating party ideologies, and the notion of culpability within the governing body. Suitably, the parliament's and government's partisan composition plays a distinctive role in decision-making processes, as it determines the general power distribution among incumbent politicians while setting the framework for joint motions, negotiations, and political opposition between the individual caucuses. In this context, the individual party ideologies can be particularly decisive as they reflect the representatives' interests and overall policy goals, while also delineating the bottom line of the opposition's stance in negotiations (Raunio and Wagner 2020, 515). Relatedly, the concept of culpability accounts for previous power constellations and for who would first be publicly blamed for any perceived wrongdoings, outlining thus which actors have to factor in policy constraints such as prohibitive costs of information sharing, or anticipated shaming campaigns. While a newly elected government for instance can blame their predecessor for their poor decision-making and encourage policy-change without any personal or partisan admission of guilt, leaders, who themselves have sanctioned contentious policy actions would most likely have to first bear high political costs for confessing their own culpability, before then being able to trigger change (Weaver 1986, 371). Hence, constraints linked to the political elite's identity shape state preferences by not only reflecting who is in power, but also by delineating the ideas the ruling politicians' stand for and the extent of perceived culpability they might have to confront upon policy-change.

In wanting to explore the applicability of these identity-based constraints, the following paper traces British and German political debates up until the year 2010. In doing so, the paper follows a timeline, which allows for an in-depth analysis of the countries' first substantial debates surrounding the introduction of anti-torture safeguards for foreigners abroad, covering thus the years between the terrorist attacks of 9/11 and the first introduction of a British safeguard. As of November 2021 Germany, has not yet publicly introduced any similar protection policy against torture abroad so that the analysis does not exclude any respective German developments. Hence, a content analysis based on three deductively derived categories of strategic constraints has been applied to 21 semi-structured interviews, which were conducted in London (February 2020) and Berlin (September 2019) with among others, different members of parliament, the former Attorney General of England and Wales, and with the former German Federal Commissioner for Human Rights Policy. Additionally, the analysis considers relevant primary sources including, for instance, parliamentary hearings, politicians' speeches, governmental reports, and press releases to filter out reoccurring themes and concepts. In this context, any strategic justification distinctively alluding to a turning point in the country's history, or the rules and culture behind a state's individual political system reflect a leitmotif of national identity, whereas any reference to a state's self-ascribed role or perceived reputation on the global stage display patterns of international identity. Finally, statements emphasizing a party's position in the government, or attempts of publicly blaming former politicians for their decisions pertain to notions of a country's political elite's identity.

The UK and Germany have been chosen as case studies as they share Western democratic values, have both faced similar pressures due to their involvement in the CIA Interrogation and Detention Program, are signatory parties to the same international regimes prohibiting torture, and yet show differing outcomes in their respective policy-making processes. The thesis focuses in particular on the states' decision-making regarding their respective intelligence services for two reasons: On one hand repeated scandals surrounding the British military's engagement in the War on Terror could construe the results, while on the other hand the German military's behavior has been less scrutinized by official investigations and parliamentary reports inhibiting thus the reliance on credible sources and the triangulation of evidence. In addition to these methodological and empirical considerations, practical implications for research and civil society likewise influenced the case selection, as so far little work has been done on the emergence of extraterritorial human rights protections, but even less work on indirect perpetrators' motivation and preferences behind such policy-making.

Empirical Analysis

As subsequently will be shown, the strategic preferences and constraints of the UK and Germany do indeed mirror country-specific patterns of identity, showcasing for instance how a state's past, international self-perception, or governmental setup impacts the state representatives' decision-making and their positioning towards human rights safeguards for foreigner abroad. Hence, in the remainder of this section, respective strategic arguments presented by British and German politicians will be dissected and scrutinized along the lines of the states' identity-based constraints, filtering out why the two states reached opposing conclusions regarding the strategic benefit of extraterritorial protections against torture.

National Constraints

United Kingdom

A key strategic argument behind the introduction of the "Consolidated Guidance" was conveying the message that the government was taking any torture-related allegations seriously and to put an end to any corresponding discussions; an endeavor especially important in the context of recent British history (Interview 7). Revelations of British collusion with the CIA provided graphic reminders of the British government's notorious and not-so-distant use of torture during the "Troubles"; thus, threatening to undermine fragile domestic relations with Northern Ireland. During the Troubles, the British military grew particularly infamous for their use of the so-called "Five Techniques", which were used as coercive interrogation methods against suspected resistance fighters who were detained by the government. Although the Five Techniques were not found to amount to torture by the European Court of Human Rights, they were nonetheless in violation of Article 3 of the European Convention on Human Rights for being inhumane and degrading; thus, resulting in their eventual prohibition in the UK in 1972 (UK Parliament 2008). As details of the UK's involvement in the CIA's detention and interrogation program became public, however, the media quickly drew comparisons to the government's history of coercive practices (Times 2009). In fact, it was eventually revealed that UK personnel had assisted in the CIA's use of EITs, including the forbidden Five Techniques, supporting accusations or a "lesson not learned" (Interview 4). This direct link between a volatile, dark period in the UK's recent history and its involvement in the War on Terror fanned the Northern Ireland's distrust in its government, and ultimately threatened to destabilize domestic relations, establishing thus a hard constraint on policy-continuance.

In addition to such concerns, British policymakers worried that change was needed to domestically protect the reputation and the morale of the British intelligence community (Interview 8). Not only was the British intelligence personnel widely recognized as a source of national pride in the UK, but there were concerns that the torture allegations would distract the MI6 from focusing on its key priority: National security in the War on Terror (Cameron 2010). Furthermore, British political culture had little tolerance for accusations against intelligence and military officers they deemed to be "heroes" (Cameron 2010). After years of hyper-analyzing British involvement in Abu Ghraib, the public grew increasingly suspicious of new allegations. Whereas the media was quick to jump on a new scandal, the British public began to criticize their government from failing to protect its intelligence officers in their fight to defend the nation; nearly 30% of the British population even indicated that in cases of emergency, torture can be justified (BBC 2006). As a result, considerable pressure was put on the government to interfere on behalf of the British intelligence agencies and to prevent any more "false" accusations by "ungrateful" and "treasonous" human rights lawyers (Interview 13).

Germany

In contrast to the UK, German history put considerably different constraints on the policymakers' strategic contemplations as no politician wanted to 'wake sleeping dogs', nor incite public indignity. It had to be expected that any policy-change could be construed as admission of guilt, which in the context of torture and against the background of WWII would carry very high political costs (Interview 18). As the German public paid rather little attention to the matter, and international accusations could be deflected with counter-references to other states' own misconduct, a narrative of denial promised more strategic advantages, especially because the opaque nature of indirect torture hampered direct comparisons with the atrocities of WWII. Furthermore, even the slightest admission of guilt could constitute a slippery slope between Germans' deeply held "never again" conviction and the people's gradually growing sensitivity towards hastily made allusions to Germany's Nazi past (Interview 15). Hence, any political motion in this regard would bare the risk of becoming the centerpiece of a very critical public debate. As a result, even policymakers supporting policy-change were constrained in their calls for policy reforms as high political costs were to be anticipated.

On a similar notion, the absence of public concern on the issue failed to justify the risk of being accused of failing to regulate the intelligence community's involvement in already contentious

military operations, a notion which given Germany's de-militarized political culture could prove to be particularly costly (Interview 21). Even if a public debate, which alluded to Germany's Nazi past, could be avoided, a policy-change indirectly indicating the previous legal extent of possible intelligence collaboration threatened to bare high political risks, as a TV documentary's mere accusation of respective CIA and BND cooperation had already triggered a significant uproar on the political level and paved the way for an in-depth parliamentary investigation (Stadler in Deutscher Bundestag 2009b, 25702). Although Germans are less inclined than Britons to view their intelligence officers as heroes, such a potential scandal would nonetheless highlight the government's failure to command its own forces according to its relatively high human rights standards. Consequently, German policymakers in their attempt to repeatedly deny Germany's complicity, were ultimately constrained in their policy-decisions, insisting ultimately that policy reforms would be redundant since German law already prohibited torture.

In sum, patterns of constraints linked to the states' national identity can be observed in both the British and German strategic reasoning. In the British case, the political-military culture, combined with the close historic ties to the Northern Ireland conflict, formed prohibitive constraints on inactivity and policy continuance. In the German case, however, the risk of policy-change being interpreted as partial admission of guilt ultimately favored a strategy of denial; especially against the background of WWII and the demilitarized German culture.

International Constraints

United Kingdom

Regarding the international realm, a key strategic drive behind Cameron's decision laid within restoring the UK's position as a role model and democratic leader in the world. Such notions were particularly rooted in the UK's self-perception as a great world power and a 'beacon of the rule of law', which not only respects, but also promotes and defends human rights around the globally (Hague 2009). Any unaddressed torture allegations would thus, not only undermine the British self-image on the international stage, but also directly contradict the state's own prioritization of global ambitions: According to which, "[British] long term interests and values are best protected by the spread of democratic values, good government, and respect for human rights" (Foreign and Commonwealth Office 2006, 35). Hence, in order to uphold the self-concept of exceptionalism and international norm-compliance, a policy-change addressing the torture allegations remained indispensable (Interview 8). As was the

case of the Consolidated Guidance and its many loopholes, however, it is important to emphasize that the presence of such subjective assessments do not necessarily translate into full norm-compliance, but can rather accumulate in patterns of window-dressing.

Another strategic constraint often linked to the policy-change in 2010, pertains to the consequences of international reputational harm for the British government caused by the pictures of Abu Ghraib and the subsequent torture-related accusations against the SIS (Interview 12). Accordingly, a country's international identity not only depends on how it positions itself of the international stage, but also on how the country's behavior is perceived by other actors (Brewster 2009, 231). By the end of the decade, the British reputation had worsened significantly among its allies: The EU openly called out the UK's complicity with the CIA, while other states questioned the reliability of the UK's intelligence community due to the court-ordered disclosure of classified operation details (Interview 10). Furthermore, David Cameron, in a speech before parliament admitted that such damage could be used by the UK's enemies to taint its reputation abroad and undermine British national security:

> For the past few years, the reputation of our security services has been overshadowed by allegations about their involvement in the treatment of detainees held by other countries. […] Our reputation as a country that believes in human rights, justice, fairness and the rule of law—indeed, much of what the services exist to protect—risks being tarnished. […] And terrorists and extremists are able to exploit those allegations for their own propaganda. (Cameron 2010, Column 175)

In the end, the British reputation and identity imposed significant constraints on the government's strategic contemplations. Without a good name, not only could future intelligence cooperation be jeopardized, but the future of the 145 British military sites abroad could also be put into question. While these circumstances alone could already threaten the country's national security, the additional risk that such details could be used to further terrorist propaganda constituted another burden to the government's security strategy.

Germany

Just like the UK, Germany likewise prides itself with being an international advocate for human rights, yet other aspects of the German self-image, notions of dependency on and loyalty to the US, seem to have particularly constrained the German decision-making on the matter (Interview 16). Hence, the strategic interest of not upsetting the main ally US constituted a recurrent theme in the debate, which was particularly often voiced in the context of ally commitments and obligations (Fischer in Zeit Online 2005). Over the years, the German

self-image as international actor has repeatedly changed, especially after WWII and the reunification in the 1990s. Throughout this time, however, the close partnership with the US constituted a stable pattern within and influence on the different areas of German policy-making, slowly evolving into a key feature of Germany's international identity (Auswärtiges Amt 2021). While in many areas this partnership translated into notions of loyalty and friendship, it also created structures of dependency, especially in the field of military and intelligence cooperation (Lübkemeier 2021, 9). When looking at the strategic calculations of German policymakers, it becomes apparent how these notions of dependency and loyalty transpired into the strategic interests of not upsetting their US counterparts, especially when any policy-change could either damage US-German relations by disclosing previous legal standards, or by undermining the future execution of respective alliance obligations.

In addition, German policymakers were likewise concerned that a new policy would introduce new hurdles for future intelligence cooperation, and thus not only tarnish its international reputation but also generate substantive strategic costs (Interview 19). Although, Germany has generally enjoyed per se a relatively good international reputation, the country has been nevertheless often criticized for band-wagoning on the efforts of other NATO members, and for not stepping up its own military presence in international conflicts (Süddeutsche Zeitung 2010b). Hence, policies, which further complicate international cooperation, could thus aggravate Germany's bureaucratic and passive reputation among its allies, making the country eventually a less attractive security partner (Interview 16). This point in particular threatened to undermine Germany's strategy for claiming international legitimacy, which is primarily anchored in multilateral approaches that require good partnerships built on solid reputations.

In the end, connections can be drawn between the British and German strategic constraints and the countries' respective international identities. Hence, Cameron's rationale favoring policy-change reflects the UK's self-perception as 'beacon of the rule of law', while also taking into account the importance of the country's international reputation. On the contrary, the German reasoning particularly emphasizes Germany's dependence on and loyalty to the US, which combined with the relatively bureaucratic and passive reputation the German military and intelligence community already has, rather discourages policy-change and additional rules.

Elite Constraints

United Kingdom

In terms of domestic political power constellations, the perspective of differentiating Cameron's own administration from the previous Labour governments promised another strategic benefit in favor of policy-change; especially as neither retribution nor increased notions of culpability had to be expected (Interview 2). The newly elected Prime Minister David Cameron and his ruling coalition between the Conservatives and the Liberal Democrats represented a new government constellation whose members had hitherto been in the opposition during the War on Terror. Given this previous outsider position in decision-making processes, the administration could take a pro-active stance instigating investigations and policy-change without fearing any accusations of culpability. In doing so, the policy reform not only promised to differentiate Cameron's government from the ones of his predecessors, but also to further emphasize the repeated failure of Tony Blair and Gordon Brown to properly address the torture-related allegations and to protect the SIS' reputation. As a result, policy-change not only had minimal transaction costs, but it also enabled Cameron's government to position itself as the more responsible and competent party in comparison to Labour.

In addition to advantages *vis-à-vis* Labour, policy reform also had the potential to aid relations within the young coalition (Interview 6). Prior to the elections held in May 2010, the Liberal Democrats clearly positioned themselves against the Iraq war, the "subservient relationship with the United States" (Liberal Democrats 2010, 63), and in general against any behavior that led to the torture allegations against the UK (Liberal Democrats 2010, 67). In their manifesto, they pledged to bring back British pride and values to the UK's foreign policy, envisioning a "full judicial inquiry into allegations of British complicity in torture and state kidnapping as part of a process to restore Britain's reputation for decency and fairness" (Liberal Democrats 2010, 68). Although Conservatives distinctly condoned the use of torture in general (Conservative Party 2010, 109), Cameron's party predominantly emphasized national security concerns in their manifesto, leaving out any allusion to the actual handling or rebutting of the reoccurring torture-allegations against the UK (Conservative Party 2010). Being perceived as compromising on the issue allowed the Conservatives not only to co-determine the wording of the policy and to thus ensure their own interest in national security, but it also constituted an opportunity to better their own position in future negotiations when they could ask the Liberal Democrats to repay the favor.

Germany

In contrast to the UK, the German government constellation had not changed since 2005, creating thus important constraints in the context of culpability and blame attribution (Interview 20). When the allegations against Germany gained traction and the parliamentary investigation published its report in June 2009, Angela Merkel had already been chancellor for four years. Hence, the likelihood that any policy change could be perceived as an admission of guilt in which her ministers could be found culpable remained relatively high (Hartmann in Deutscher Bundestag 2009, 25705); given the pending federal elections in 2009, the costs of such accusations seemed particularly disadvantageous. Furthermore, at the time of the debate, Germany was ruled by the so-called "Big Coalition" („*GroKo*") between CDU/CSU and the SPD; therefore, any in-depth investigations into allegations against former Chancellor Schroeder's (1989-2005) SPD government, could have severe consequences for Merkel's coalition partner. A scandal in which key politicians from either party were linked to misconduct would not only be disastrous for their individual party, but it could also undermine the future of the entire ruling coalition. In light of these considerations, strategic constraints associated with the introduction of policy reforms emerged for both ruling parties, especially as there was minimal public pressure to do otherwise.

Although both parties recognized the importance of human rights and the ban on torture, both CDU/CSU and SPD ran on campaigns that prioritized national security matters (Interview 15). In fact, several high-ranking members of both parties publicly supported the use of torture in "ticking-time bomb scenarios," thus illustrating the parties' perspective on the matter. In an interview, then Social Democrat Minister of Justice Brigitte Zypries, for instance, reiterated the importance of the ban on torture, but added that German law allowed for extraordinary measures in justified emergency situations (Deutsche Welle 2003). In a similar fashion, then Minister of Internal Affairs Wolfgang Schäuble (CDU) stated that Germany stands by the prohibition of torture, but that in matters of preventing terrorism it would be irresponsible not to use evidence, only because it was unclear whether or not foreign intelligence services had gained the information through torture (Süddeutsche Zeitung 2010a). Hence, in light of their party priorities, any additional restrictions could be seen as undermining national security priorities, and thus, too costly. Consequently, German policymakers preferred upholding their narrative of Germany "does not torture," which not only preserved operational flexibility, but also was relatively hard to contest, given the opaque nature of the complicit behavior.

Ultimately, both countries elite's identity played a decisive factor in constraining their respective decision-making. While Cameron's newly elected government could pass the bucket to previous administrations, Angela Merkel and her coalition government faced high political costs in case of a policy-change being interpreted as admission of guilt. Similarly, German decision makers valued human rights, but ultimately prioritized national security, whereas Cameron lead a coalition where one partner insistently sought for in-depth investigations into the torture-allegations.

Conclusion

This article investigates why in 2009 and 2010, despite very similar contextual parameters and attempts of being held accountable, the British and German governments came to diametrical different conclusions when deciding about the introduction of anti-torture safeguards for foreigners abroad. In doing so, the paper first briefly scrutinizes each country's strategic rationale behind their decision-making, before then translating insights from identity literature to the analysis of constraints, by investigating how the identity-based limitations pertaining to the national, international, and political elite's level have shaped the government's preferences and strategic contemplations. In both cases, a connection between these multi-layered constraints and the representatives' final decision-making can be drawn enabling thus, a more nuanced and refined understanding for the differing state preferences and policy-decisions. Accordingly, close historic ties, a jeopardized international reputation, and little risk of culpability further motivated Cameron's decision of policy-change, whereas the same country-specific parameters in Germany paved exactly the opposite path.

Due to its many loopholes and inconsistencies, Cameron's policy output can hardly be called a panacea against torture, but rather a window-dressing measure (see Hafner-Burton and Tsutsui 2005, 1378). Nonetheless, this article still shows how considering the multiple layers of a state' strategic preferences can render important information to future campaigns of social influence (see Murdie and Davis 2012; Donno and Neureiter 2018). For instance, with a solid understanding for the decision-makers' country-specific, multi-faceted constraints, policy entrepreneurs can tailor their strategic interventions, negotiation efforts, and learning processes to the countries vulnerabilities rather than pursuing a generic one-size-fits-all approach. Hence, while allusions to British historic responsibilities seem to have motivated change, similar approaches in Germany have rather solidified the decision-makers' determination for policy-continuance. If understanding, however, the emphasis German

decision-makers put on being perceived as loyal cooperation partner to the US, future accountability campaigns could benefit from rather targeting these relations than trying to cause uproar among the German public, which in the last twenty years has mostly remained silent or indifferent on the issue.

In line with this, the research shows how targeted shaming as an accountability mechanism for indirect human rights violations is more effective if it focuses on the long-term consequences of undermining a state's self-image on the international stage, rather than solely trying to trigger immediate public uproar (see Sikkink 2013). Accordingly, both states were at different points of time subjected to public shaming campaigns, however, both states withstood the pressure by engaging in different evasion techniques (see e.g. Blakely and Raphael 2020) resulting ultimately in neither Merkel, nor Tony Blair nor Gordon Brown pursuing policy change. It was not until Cameron was motivated on one hand by the new coalition partner, but more importantly by the credible, anticipated threat to the British international identity that he eventually introduced the new guidance to avoid such a scenario. Despite such a strategic rather than normative proceeding, and regardless of the many loopholes Cameron's policy entailed, it still increased the transparency of British intelligence operations; an important step, which never took place in Germany where the shaming campaigns, not taking into account the particularities of German identities, failed to create a credible threat to the country's carefully crafted (inter)national image.

In the end, there remain three avenues for future research. First, while this analysis did translate insights from identity literature into a more nuanced understanding of strategic constraints, it did so in the context of a rationalist analysis, rather than pursuing a full-fledged Constructivist examination. While such a proceeding offered important insights and inferences, Constructivist assessments regarding the primacy of individual identities (see Epstein 2011, Hopf and Allan 2016) could shed further light on the ranking of preferences and the hierarchy of respective constraints. Secondly, future studies looking into additional cross-case comparisons could generate a greater generalizability of the findings and filter out which identity patterns are particularly prone to trigger decisions in favor of policy-change, and which patterns rather encourage policy-continuance. In this context, it might also be of interest to conduct comparisons with non-Western accomplices of the CIA program, like Thailand, to see if the importance of identity patterns also transgresses culture boundaries. Finally, another avenue adding to a greater generalizability would be to test the interplay between identity patterns and strategic decision-making on additional cases of non-torture related human rights

violations, in order to further explore the states' motivation behind blocking or introducing extraterritorial human right safeguards.

Bibliography

Adler, Emanuel. 2013. "Constructivism in International Relations: Sources, contributions, and debates." In *Handbook of international Relations* by Thomas Risse; Beth A. Simmons and Walter Carlsnaes, 112-144. London: Sage Publication.

Adler Emanuel, and Vincent Pouliot. 2011. "International practices." *International Theory* 3, no.1: 1-36.

Alden, Chris. 2017. "Critiques of the rational actor model and foreign policy decision making." In: *Oxford Research Encyclopedia of Politics* by Thompson, William R. and Capelos, Tereza. Oxford: Oxford University Press.

Amnesty International. 2009. "Gegen den Strom." Accessed 15 September, 2021. https://www.amnesty.de/journal/2009/juni/gegen-den-strom.

Anckar, Carsten. 2008. "On the applicability of the most similar systems design and the most different systems design in comparative research." *International Journal of Social Research Methodology* 11, no.5: 389-401.

Auswärtiges Amt. 2021. "United States of America: bilateral relations to Germany." Accessed September 04, 2021. https://www.auswaertiges-amt.de/en/aussenpolitik/unitedstatesofamerica/218718.

Baldwin, David. 2012. "Power and international." In *Handbook of international Relations* by Thomas Risse; Beth A. Simmons and Walter Carlsnaes, 273-297. London: Sage Publication.

BBC. 2006. "One-third support some torture." Accessed 06 September, 2021. http://news.bbc.co.uk/2/hi/in_depth/6063386.stm.

Blakeley, Ruth. and Raphael, Sam. 2017. "British torture in the 'war on terror." *European Journal of International Relations* 23, no.2: 243-266.

Blakeley, Ruth, & Raphael, Sam. 2020. Accountability, denial and the future-proofing of British torture. *International Affairs* 96, no.3: 691-709.

Hawkins, Darren. 2004. "Explaining costly international institutions: Persuasion and enforceable human rights norms." *International Studies Quarterly* 48, no.4: 779-804.

Barnett, Michael. 2018. "Constructivism." In *The Oxford Handbook for International Security* by Alexandra Gheciu, and William C. Wohlforth, 86-100. Oxford: Oxford University Press.

Brewster, Rachel. 2009. "Unpacking the State's Reputation." *Harvard International Law Journal* 50: 231-270.

Bundesgesetzblatt. 2009. "Gesetz zur Fortentwicklung der parlamentarischen Kontrolle der Nachrichtendienste des Bundes." Accessed August 04, 2021. https://www.bgbl.de/xaver/bgbl/start.xav#__bgbl__%2F%2F*%5B%40attr_id%3D%27bgbl109s2346.pdf%27%5D__1628065614723.

Bundesministerium der Verteidigung. 2021. "Vernetzter Ansatz." Accessed September 04, 2021. https://www.bmvg.de/de/themen/friedenssicherung/vernetzter-ansatz.

Cameron, David. 2010. "Treatment of Detainees." Accessed September 01, 2021. https://hansard.parliament.uk/commons/2010-07-06/debates/10070631000002/TreatmentOfDetainees.

Conservative Party. 2010. "An Invitation to Join the Government of Britain – The Conservative Manifesto 2010." Accessed September 03, 2021. https://conservativehome.blogs.com/files/conservative-manifesto-2010.pdf.

Creamer, Cosette D., and Beth A. Simmons. 2019. "Do self-reporting regimes matter? Evidence from the convention against torture." *International Studies Quarterly* 63, no.4: 1051-1064.

Deutscher Bundestag. 2009a. "Beschlussempfehlung und Bericht des 1. Untersuchungsausschusses nach Artikel 44 des Grundgesetzes 18 June 2009." Accessed July 28, 2021. https://dserver.bundestag.de/btd/16/134/1613400.pdf.

Deutscher Bundestag. 2009b. "Deutscher Bundestag 230. Sitzung 02. Juli, 2009." Accessed August 03, 2021. https://dserver.bundestag.de/btp/16/16230.pdf#P.25700.

Deutsche Welle. 2003. "Police Threat Fuels Debate On Torture." Accessed September 03, 2021). https://www.dw.com/en/police-threat-fuels-debate-on-torture/a-785751.

Deutsche Welle. 2005. "Schäuble: Facts Obtained Through Torture Should be Used." Accessed August 03, 2021. https://www.dw.com/en/sch%C3%A4uble-facts-obtained-through-torture-should-be-used/a-1822384.

Donno Daniela and Neureiter Michael. 2018. "Can human rights conditionality reduce repression? Examining the European Union's economic agreements." *The Review of International Organizations* 13, no.3: 335-357.

Duffield, John. 1998. "*World power forsaken: political culture, international institutions, and German security policy after unification.*" Palo Alto: Stanford University Press.

Epstein, Charlotte. 2011. "Who speaks? Discourse, the subject and the study of identity in international politics." *European Journal of International Relations* 17, no.2: 327-350.

Epstein, Charlotte. 2013. "Theorizing agency in Hobbes's wake: the rational actor, the self, or the speaking subject?" *International Organization* 67, no.2: 287-316.

European Parliament. 2006. "Alleged secret detentions and unlawful inter-state transfers of detainees involving Council of Europe member states." Accessed March 30, 2021. https://assembly.coe.int/CommitteeDocs/2006/20060606_Ejdoc162006PartII-FINAL.pdf.

European Parliament. 2007. "Report on the alleged use of European countries by the CIA for the transportation and illegal detention of prisoners (2006/2200(INI))." Accessed August 03, 2021).https://www.europarl.europa.eu/comparl/tempcom/tdip/final_report_en.pdf.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49, no.3: 379-414.

Fearon, James, and Alexander Wendt. 2002. "Rationalism v. Constructivism: a skeptical view." In *Handbook of International Relations* by Thomas Risse; Beth A. Simmons and Walter Carlsnaes, 52-72. London: Sage Publications.

Foreign and Commonwealth Office. 2006. "Active Diplomacy for a Changing World." Accessed September 01, 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/272260/6762.pdf.

Frankfurter Allgemeine Zeitung. 2007. "Fall Kurnaz : Neue Vorwürfe: Steinmeier in Erklärungsnot." January 20, 2007. https://www.faz.net/aktuell/politik/fall-kurnaz-neue-vorwuerfe-steinmeier-in-erklaerungsnot-1412129.html.

Frieden, Jeffrey A. 1999. "Actors and preferences in international relations." In: *Strategic Choice and International Relations* by Lake David A. and Powell, Robert, 39-76. Princeton: Princeton University Press.

Gillis, John R., ed. 1994. "*Commemorations: The politics of national identity*." Princeton: Princeton University Press.

Hafner-Burton, Emille M. and Tsutsui, Kiyoteru 2005. "Human rights in a globalizing world: The paradox of empty promises." *American Journal of Sociology* 110, no.5: 1373-1411.

Hafner-Burton, Emilie M. and Shapiro, Jacob N. 2010. "Tortured relations: Human rights abuses and counterterrorism cooperation." *PS: Political Science & Politics* 43, no.3: 415-419.

Hafner-Burton, Emilie M., Haggard Stephan, Lake, David A., and David G. Victor. 2017. "The behavioral revolution and international relations." *International Organization* 71, no.1: 1-31.

Hague, William. 2009. "Remarks at British Conservative Party Conference." Accessed 04 September, 2021. https://www.c-span.org/video/?289357-2/william-hague-remarks-british-conservative-party-conference.

HM Government. 2010. "Consolidated Guidance to Intelligence Officers and Service Personnel on the Detention and Interviewing of Detainees Overseas, and on the Passing and Receipt of Intelligence Relating to Detainees." Accessed March 28, 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/62632/Consolidated_Guidance_November_2011.pdf.

HM Government. 2019. "The Principles relating to the detention and interviewing of detainees overseas and the passing and receipt of intelligence relating to detainees." Accessed March 28, 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/818306/20190718_The_Principles_relating_to_the_detention_and_interviewing_of_detainees_overseas.pdf.

Hopf, Ted. 2002. "*Social construction of international politics: Identities & foreign policies, Moscow, 1955 and 1999.*" Ithaca: Cornell University Press.

Hopf, Ted and Allan, Bentley (Eds.). 2016. *Making identity count: building a national identity database*. New York: Oxford University Press.

House of Commons. 2010. "Oral Answers to Questions Foreign and Commonwealth Office." Accessed August 02, 2021. https://publications.parliament.uk/pa/cm201011/cmhansrd/cm100706/debtext/100706-0001.htm.

Human Rights Watch. 2006. "Undermining the Torture Ban." Accessed March 16, 2022. https://www.hrw.org/legacy/backgrounder/eca/uk1106/4.htm.

Hurd, Ian .1999. "Legitimacy and authority in international politics." *International Organization* 53, no.2: 379-408.

Intelligence and Security Committee. 2015. "The Handling of Detainees by UK Intelligence Personnel in Afghanistan, Guantanamo Bay and Iraq." Accessed August 02, 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/224690/isc_terrorist_attacks_7july_report.pdf.

Intelligence and Security Committee. 2018. "Detainee Mistreatment and Rendition: 2001–2010." Accessed July 28, 2021. https://fas.org/irp/world/uk/isc-detainee.pdf.

Jupille, Joseph, Caporaso, James A., and Jeffery T. Checkel. 2003. "Integrating institutions: Rationalism, constructivism, and the study of the European Union." *Comparative Political Studies* 36, no.1-2: 7-40.

Johansen, Stian Ø. 2020. "*The Human Rights Accountability Mechanisms of International Organizations*." Cambridge: Cambridge University Press.

Liberal Democrats. 2010. "Liberal Democrat Manifesto 2010." Accessed September 03, 2021. https://www.markpack.org.uk/files/2015/01/Liberal-Democrat-manifesto-2010.pdf.

Liese, Andrea. 2009. "Exceptional necessity- How liberal democracies contest the prohibition of torture and ill-treatment when countering terrorism." *Journal of International Law and International Relations 5,* no.1: 14-47.

Luban, David. 2007. "Liberalism, torture, and the ticking bomb." In *Intervention, Terrorism, and Torture* by Lee, Steven P., 249-262. Dordrecht: Springer.

Lübkemeier, Eckhard. 2021. "Standing on Our Own Feet? Opportunities and Risks of European Self-Defence. SWP Research Paper." Accessed September 04, 2021. https://www.swp-berlin.org/publications/products/research_papers/2021RP01_EuropeanSelfDefence.pdf.

Manners, Ian, and Richard Whitman. 2003. "The 'difference engine': constructing and representing the international identity of the European Union." *Journal of European Public Policy* 10, no.3: 380-384.

Murdie, Amanda M. and Davis, David R. 2012. "Shaming and blaming: Using events data to assess the impact of human rights INGOs." *International Studies Quarterly* 56, no.1: 1-16.

Nowak, Manfred. 2010. "Folter und die Verwertung von Informationen bei der Terrorismusbekämpfung." Accessed September 05, 2021. https://www.ecchr.eu/fileadmin/Publikationen/Folter_und_die_Verwertung.pdf.

Pye, Lucian W. 2015. "Introduction: Political Culture and Political Development." In *Political Culture and Political Development* by Lucian W. Pye, and Sidney Verba, 1-26. Princeton: Princeton University Press.

Raunio, Tapio, and Wolfgang Wagner. 2020. "The Party Politics of Foreign and Security Policy." *Foreign Policy Analysis* 16, no.4: 515-531.

Ryan, Lorraine. 2012. "Cosmopolitan memory and national memory conflicts: On the dynamics of their interaction". *Journal of Sociology* 50, no.4: 501–514.

Scharpf, Fritz W. 2018. "*Games real actors play: Actor-centered institutionalism in policy research*." New York: Routledge.

Shannon, Brooke N., McGee, Zachary A., and Jones, Bryan D. 2019. "Bounded rationality and cognitive limits in political decision making." *Oxford Research Encyclopedia of Politics*. Accessed September 06, 2021. https://oxfordre.com/politics/view/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-961.

Sikkink, Kathryn. 2013. "The United States and torture: Does the spiral model work?." *Cambridge Studies in International Relations* 126, no.1: 145-163.

Simon, Herbert A. 1997. "*Models of bounded rationality: Empirically grounded economic reason"* (Vol. 3). Cambridge: MIT Press.

Snidal, Duncan. 2012. "Rational choice and international relations." In *Handbook of International Relations* by Thomas Risse; Beth A. Simmons and Walter Carlsnaes, 85-111. London: Sage Publications.

Süddeutsche Zeitung. 2010a. "CDU-Ministerpräsident kritisiert Schäuble." Accessed September 03, 2021. https://www.sueddeutsche.de/politik/verwertung-von-folter-gestaendnissen-cdu-ministerpraesident-kritisiert-schaeuble-1.653331.

Süddeutsche Zeitung. 2010b. "Kriegsscheue Deutsche Trittbrettfahrer." Accessed September 03, 2021. https://www.sueddeutsche.de/politik/bundeswehr-einsatz-in-afghanistan-kriegsscheue-deutsche-trittbrettfahrer-1.293070.

Tannenwald, Nina. 2005. "Ideas and explanation: Advancing the theoretical agenda." *Journal of Cold War Studies* 7, no.2: 13–42.

The Rendition Project. 2022. "Rendition Circuits." Accessed March 15, 2022. https://www.therenditionproject.org.uk/flights/renditions/index.html.

Times. 2009. "Techniques of the Troubles outlawed in 1972." Accessed September 04, 2021. https://www.thetimes.co.uk/article/techniques-of-the-troubles-outlawed-in-1972-rzpgkffhh6r.

UK Parliament. 2008. "Joint Committee on Human Rights Twenty-Eighth Report." Accessed September 04, 2021. https://publications.parliament.uk/pa/jt200708/jtselect/jtrights/157/15703.htm.

Weaver, R. Kent. 1986. "The politics of blame avoidance." *Journal of Public Policy* 6, no.4: 371-398.

Weedon, Chris, and Glenn Jordan. 2012. "Collective memory: theory and politics." *Social Semiotics* 22, no.2: 143-153.

Wendt, Alexander. 1994. "Collective identity formation and the international state." *American Political Science Review* 88, no.2: 384-396.

Zeit Online. 2005. "Bedingt aussagebreit". Accessed September 05, 2021. https://www.zeit.de/2005/52/Fischer?utm_referrer=https%3A%2F%2Fwww.google.com%2F.

Interviews

Interview 1: Former Senior Member UK National Preventive Mechanism, London 11/02/2020.

Interview 2: Member of Parliament, Conservative Party, London 04/02/2020.

Interview 3: Member of Parliament, Labour Party, London 12/02/2020.

Interview 4: Senior Lecturer in Politics and International Relations, London 14/02/2020.

Interview 5: Former Senior Member of the Equality and Human Rights Commission, London 17/02/2020.

Interview 6: Spokesperson Reprieve, London, 10/02/2020.

Interview 7: Former Member Gibson Inquiry, London, 25/02/2020.

Interview 8: Former Member United Nations Subcommittee for the Prevention of Torture, London, 25/02/2020.

Interview 9: Former Security and Intelligence Coordinator in the Cabinet Office, London, 04/03/2020.

Interview 10: Professor of Foreign Policy and International Relations, London, 05/03/2020.

Interview 11: Former Barrista at Public Interest Lawyers Ltd., Skype, 09/04/2020.

Interview 12: Former Attorney General of England and Wales, Skype, 27/04/2020.

Interview 13: Reader in International Law, Birmingham, 17/02/2020.

Interview 14: Spokesperson Human Rights Watch, Skype, 13/08/2019.

Interview 15: Former Director German Institute for Human Rights, Skype, 26/09/2019.

Interview 16: Former Federal Commissioner for Human Rights Policy and Humanitarian Aid, Berlin, 20/09/2019.

Interview 17: Former Member of Parliament and Member of the BND Untersuchungsausschuss 2009, die Linke, Skype, 10/09/2019.

Interview 18: Member of Parliament, FDP, Berlin, 12/09/2019.

Interview 19: Professor for International Law, Berlin, 09/09/2019.

Interview 20: Spokesperson Aktion der Christen für die Abschaffung der Folter, Skype, 14/11/2019.

Interview 21: Professor for International Relations, Berlin, 09/09/2019.

<div align="center">**Article 3:**</div>

<div align="center">**British counterterrorism, the international prohibition of torture, and the Multiple Streams Framework**</div>

Abstract

After years of violating the basic principles of human rights in the name of counterterrorism, Western democracies have begun to implement extraterritorial safeguards that extend protections under the Convention against Torture to foreigners abroad. The case of the UK and the development of the 'Principles' in 2019, however, presents a particular puzzle to policymaking research, as it challenges traditional hypotheses regarding the opening of problem windows within the multiple streams framework. Accordingly, the UK presents an interesting case in which a powerful state willingly engaged in self-restraint, despite little electoral pressure to do so and a persistently high terrorist threat. Drawing on theory-building process-tracing, this article addresses this gap using data from semi-structured interviews with British policy experts to present a refined hypothesis, which can also be applied to policy fields of little public interest and processes of foreign policymaking.

Keywords: Foreign policy; human rights; counterterrorism; multiple streams framework; convention against torture; UK

1. Introduction

"The more a condition puts the policy makers' re-election at risk, the more likely it is to open a policy window in the problem stream" (H1) (Herweg et al., 2015: 437) – this hypothesis introduced by Herweg, Huß, and Zohlnhöfer remains widely accepted in the Multiple Streams Framework (MSF), but while it offers important insights into the nexus of domestic pressures and policy-making, it simultaneously raises a different question: How can we explain the opening of problem windows, when the triggering conditions do *not* put the decision-makers' re-election at risk? Take, for example, the development of British anti-torture safeguards by Theresa May's Administration in the midst of Brexit. By providing pivotal intelligence, questions sets, and logistic support for rendition flights, the British Secret Intelligence Service (SIS) not only colluded with the American Central Intelligence Agency (CIA), but also benefitted from information gained through torturous practices (ISC, 2018a: 3). Despite the publication of a parliamentary report, which deemed British anti-torture safeguards as insufficient and acknowledged the UK's complicity in the American post-9/11 Detention and Interrogation Program (ISC 2018a&b), the British government faced little pressure to initiate

any reform. In fact, the "Principles"[39] were passed amidst open resistance from the public to any further investigations into misconduct by the military and intelligence agencies, during which human rights lawyers were frequently condemned in political debates as left-wing activists, who "harangue and harass the bravest of the brave [the British military]" (May,2016). Nevertheless, May, whose own political interests, ambitions, and priorities rather contradicted any push for further extraterritorial human rights protections (ibid.), immediately acted upon the parliament's negative feedback, tasking the Investigatory Powers Commissioner's Office[40] (IPCO) to come up with a fitting policy-solution to the problem. Even though the final policy-output does not constitute a panacea against torture, this case illustrates a limitation in the prominent MSF literature, particularly relating to problem windows opening in issue areas of little public interest and foreign policy-making processes, providing thus an opportunity to refine Herweg et al.'s original hypothesis.

When looking at MSF literature, problem windows are well discussed (see e.g. Knaggard, 2015; Béland, 2016; Zahariadis, 2016), despite this central role of the concept, however, current literature remains fairly vague and fails to comprehensively explain when, in absence of electoral pressure, a condition passes the threshold of opening a problem window and how resulting coupling processes unfold (see Travis and Zahariadis, 2002: 496). Other policy models, such as Historical Institutionalism or the Punctuated Equilibrium Theory, might use different terminology and paradigms, but ultimately scrutinize the appearance of problems and their impact on sudden policy-changes as well. By analyzing the concept of critical junctures, scholars of Historical Institutionalism, for instance, provide extensive insights on how sudden problems materialize and how they subsequently influence the hitherto chosen policy-paths of governments (Capoccia and Kelemen, 2007). Yet, while such notions can help detect whether or not a problem did actually pass the threshold of the problem agenda, they nonetheless fail to shed further light on why they did so in first place.

In light of this gap, this article follows a theory-building process tracing approach to inductively derive new theoretical insights from the empirical data pertaining to the British case study (Beach & Pedersen, 2019). As will be shown in this study, the making of the British "Principles" is, in contrast to other traditional (*doctrinal coupling,* see Zahariadis, 2003: 72**)** as well as new (*spillover commissioning,* see Ackrill & Kay, 2011: 77) MSF mechanisms, best explained by the concept of *consequential coupling*; though above mentioned incongruities require a further refinement of the hypothesis postulated by Herweg et al. Hence, the study adheres to the general logic of the Multiple Streams Framework (MSF), but remains open to inductively gained inferences regarding the opening of problem windows in absence of electoral pressures. In order

to do so, the article relies on a multitude of primary sources as well as on 13 interviews conducted with among others British Members of Parliament, former Executive employees, and the former Attorney General for England and Wales.

As a result, the following research addresses the limits of current MSF theory by proposing a refined hypothesis on the opening of problem windows, which can be applied to both domestic as well as foreign policy-making processes as it de-emphasizes the primacy of electoral pressures without losing the essence of Herweg et al.'s original hypothesis. Accordingly, the article argues that *the more a condition puts the policymakers' influence on negotiations and general decision-making processes at risk, the more likely it is to open a policy window in the problem stream* (H2). Regarding domestic policy issues, this influence is endangered if a policymakers' re-election, and thus their position at the domestic negotiation table is at risk. Regarding foreign policy issues, the policymakers' influence is endangered if their personal credibility as state representative is at risk or if the problem jeopardizes the state's general reputation and thus its negotiating power in the international order.

In light of these results, the article contributes to the academic debate in three ways: It enhances the theoretical understanding of policy-making processes, provides new empirical evidence, and bridges the field of policy research with the field of international human rights compliance. To begin with, the article inductively refines Herweg et al.'s (2015) hypothesis on the opening of problem windows, which not only broadens the MSF's understanding for processes where politicians act against the public will and their own political program, but also opens the theoretical model to foreign policy analyses. The thorough analysis of new interview material and primary sources adds to the existing literature by offering a comprehensive picture of the events leading-up to the British "Principles" in 2019. Finally, the study applies a traditional policy-making theory to a topic that is usually located in the field of IR or international law, connecting thus research areas, which hitherto have mostly been kept separated.

The article's first section outlines the research design by delving into the specifics of theory-building process tracing, the criteria for case selection, and further details regarding data collection. Afterwards, section two offers an in-depth analysis of the case study, the respective problem window and coupling process, before the third section presents the newly refined hypothesis and discusses further interesting findings such as Theresa May's pro-active role in, as well as Brexit's influence on the making of the "Principles".

2. Research Design

Theory-building process tracing entails the analysis of an empirical narrative, which diverges from prevailing theoretical assumptions, and therefore enables the refinement of prevalent hypotheses and mechanisms **(Geddes, 2003).** "The goal is to translate this descriptive narrative into a blow-by-blow account that can plausibly link together C[ause] and O[utcome] as we attempt to figure out key causal steps in the process" (Beach & Pedersen, 2019: 277). Even though such inductive approaches can sometimes uncover completely new theoretical schools of thought, it is still more common that respective research is inspired by an already existing theory as crucial components and causal linkages might differ, but the theory's overall rationale remains applicable (Geddes, 2003). Given the mismatch between the empirical data and Herweg et al.'s hypothesis on the opening of problem windows, the following article draws on a vast range of primary and secondary sources to identify systematic patterns in the empirics and to explain the events from an MSF perspective.

The study focuses on one particular case, the British policy-making process behind the adoption of the "Principles" (HM Government, 2019). The case selection is based on three considerations: First, the interplay of the parliamentary reports' publication [Cause, "C"] and the subsequent elaboration of a new extraterritorial safeguard [Outcome, "O"] suggest the existence of an underlying causal mechanism, which links C and O together and facilitates a theory-building process tracing approach (Rohlfing, 2014: 617). Second, the empirical data tracing the development of the British safeguard indicates various incongruities with existing MSF mechanisms, posing a theoretical puzzle. Finally, aside from these theoretical and methodological considerations, the UK also poses an interesting empirical case given the state's active involvement in global counterterrorism operations, and May's pro-active push for policy-change despite the public's indifferent if not opposing stance to the matter.

Despite their limited generalizability, single-case study designs have become common practice in inductive research approaches as such analyses provide detailed insight into the process in questions, and thus, allow for well-informed inferences and theoretical postulations (Sikkink, 2013). When looking at a single case, the time and means that would otherwise be divided across multiple cases can be used to focus extensively on one specific process (Gerring, 2004), even if that means sacrificing cross-case comparisons. This concentration of resources, in turn, enables the researcher to map out complex developments in detail and to identify pivotal policy events and paths that might have influenced the process' deviation from the existing theoretical models (Siggelkow, 2007). Correspondingly, theory-building process tracing is subsequently

used to draw clear-cut causal inferences, which afterwards, in future research, can be empirically tested via a "snowballing-outward strategy" to generate a higher level of generalizability (Beach & Pedersen, 2019: 278).

In order to meet the standards of in-depth process tracing, the analysis is based on empirical data gathered from primary and secondary sources as well as on interviews conducted with British decision-makers and stakeholders. Primary sources including politicians' public speeches, statements, and meeting minutes have been analyzed to filter out the decision-makers' motivation for policy-change; whereas, secondary sources were key to understanding interrogation guidelines, identifying central actors in the respective policy-making processes, and grasping current trends in the British anti-torture debate. In addition to the analysis of primary and secondary sources, interviews with key actors have been crucial in gaining further information, detailed insights, and first-hand accounts. Hence, in spring 2020, almost ten hours of recorded material has been gathered in 13 semi-structured interviews, which have been conducted remotely as well as personally in London and Birmingham; among others with the former General Attorney of England and Wales, Members of Parliament, and representatives from the Executive and civil society. The interviews used throughout this research follow the ethical guidelines of qualitative research (Miller et al., 2012) so that all interviewees have been fully informed about the purpose and the object of this study, while furthermore their privacy and confidentiality remain respected. All interview partners have been approached via email.

3. The Multiple Streams Framework

Theoretical Embedding

Since John W. Kingdon (1984) first introduced the Multiple Streams Framework, the model quickly grew popular for its ability to explain why at a certain point in time, some policy proposals make it onto the decision-agenda, while other proposals simply disappear (Van der Heijden et al., 2019: 5). The theory maintains that policy-making procedures are not inherently rational and linear processes, but rather the product of contingent stream development and a policy entrepreneur's successful exploitation of a policy window (Herweg et al. 2015) following the successful alignment of three streams (Kingdon 1984: 19). The problem stream describes how policy entrepreneurs or problem brokers exploit focusing events, indicators, or negative feedback[41] in order to gain the politicians' and/or public's attention, and to compel the incumbent decision-makers to take action (Brunner, 2008: 52). The political stream, in contrast, describes how the general public mood, electoral turnovers, or a politician's ideology can

influence prevailing political dynamics and alter the decision-makers' disposition for change (Cairney & Zahariadis, 2016: 93). While in these two streams a policy window, a fleeting opportunity for advocates to push for policy-change, can open (Farley et al., 2007), the policy stream rather focuses on the development of feasible policy alternatives. Accordingly, ideas emerge within the respective policy communities, where they are evaluated, modified, and combined until they are narrowed down to a short-list of feasible policy options (Zhu, 2008: 317). In order for policy-change to occur, a policy entrepreneur must invest their resources and skills to successfully align all three streams. In MSF, lobbyists, stakeholders, and academics usually take on the roles of policy entrepreneurs, but policy experts within the different parties (Herweg et al., 2015), or elected officials themselves (Zohlnhöfer, 2016: 89) can also advocate for policy-change.

In light of this, and based on the assumption of ambiguity being a core component of any policy-making processes (Copeland & James, 2014: 2), MSF foresees three causal mechanisms that all share the same cause and outcome, but ultimately pertain to different constellations of the above-mentioned building blocks and alignment processes (see Fig. 1). Hence, all three causal mechanisms start with patterns of independent, contingent stream development, as the discrepancy between short-lived problem attention and the long process of solution finding renders scenarios of linear policy-making processes as rather unlikely (Cairney and Zahariadis, 2016: 87). The mechanisms' common end, however, is the final policy adoption, which, for instance, enhances the statute's scope, increases the level of accountability, provides more precise legal language, or expands the legal bindingness of an already existing safeguard (Fariss: 2014: 299; Abbott et al. 2000: 408-9). In assuming policy adoption as final point, the study differs from Kingdon's original theory, which solely focuses on the specification of alternatives. Yet, in light of the Executive nature of counter-terrorism reforms, such a convergence of agenda-setting and policy adoption seems plausible as related governmental policy-changes do not require further rounds of negotiation, nor additional parliamentary approval processes. In policy-fields that do however necessitate additional parliamentary adoption procedures, further analyses of the corresponding mediation and decision-making cycles should be considered and added to the model (see Herweg et al., 2015).

Cause	→	Policy Window	→	Stream Alignment	→	Outcome
Doctrinal Coupling						
Contingent Stream Development	→	Window opens in the Political Stream enabling politician(s) to promote their pet policy	→	Politician(s) engages in problem-oriented alliance-building	→	Policy adoption
Consequential Coupling						
Contingent Stream Development	→	Window opens in the Problem Stream compelling politician(s) to find a quick solution	→	Policy entrepreneur(s) targets politician(s) for policy-selling	→	Policy adoption
Spillover Commissioning						
Contingent Stream Development	→	Window opens in the Political Stream forcing politician(s) to manage institutional spillovers	→	Politician(s) selects already known policy proposal	→	Policy adoption

Figure 1: Causal Mechanisms of Stream Alignment

The first mechanism, *doctrinal coupling*, accounts for processes in which an elected official adapts their political agenda and starts pushing for a new pet policy, after changing political dynamics have opened a policy window in the politics stream (Herweg et al., 2018: 27). As the official's policy-seeking motivation is often not transferable to other decision-makers and consequently hampers corresponding persuasion and bargaining efforts, the elected official, acting as a political entrepreneur, evokes a hitherto unrelated, but fitting narrative from the problem stream to substantiate and broaden their policy claim (Boscarino, 2009: 416). Once the majority of politicians supports the envisioned reform, the decision-agenda's amendment and the subsequent policy adoption ensue (Zahariadis, 2003: 72). *Consequential coupling*, in contrast, describes how policy entrepreneurs frame a focusing event, relevant indicators, or negative feedback to convince decision-makers of adapting their problem agenda (ibid.). If successful, they can then step in to sell their proposal to the elected officials by presenting their ideas as best option available (Jones, 2003: 396). Lastly, *spillover commissioning*, outlines how institutional spillovers from one policy field to another open a policy window in the politics stream by putting the affected officials under such a reform pressure that they actively seek the help of the policy community (Ackrill & Kay, 2011). "In the search for legitimate solutions […], policy makers will tend to choose from an existing menu of policies in order to maximise

value acceptability and technical feasibility" (Copeland & James, 2014: 8). Upon finding a fitting, pre-existing policy alternative, the elected official amends the decision-agenda to the seemingly new policy (Ackrill & Kay, 2011).

4. The Making of the "Principles"

As will be subsequently shown, the policy-making process behind the British "Principles" follows the logic of *consequential coupling*, but diverges from previous assumptions, as even in absence of related electoral pressures, the parliament's negative feedback still triggered the opening of a problem window. The remainder of the section will delve into an in-depth analysis of the British case study, unraveling the emergence of the problem window, demonstrating the subsequent unfolding of the consequential coupling process, and illustrating the need to refine Herweg et al.'s hypothesis in order to account for developments on the international stage.

Contingent Stream Development

Upon entering the 'War on Terror', the UK not only provided military help to the US, but also warranted the British secret intelligence services' (SIS) assistance to the American Central Intelligence Service (CIA) and its Detention and Interrogation Program (European Parliament, 2007: Sec. 67-78). In consequence, the UK was found complicit in severe breaches of the CAT, as British personnel provided on at least 232 occasions pivotal information and question sets to their American counterparts "despite knowing or suspecting that the detainee had been or was being mistreated" (ISC, 2019: 3). British officers also reportedly left an ongoing interrogation, only to return after the detainee had been "roughed up" and showed signs of being "physically hurt" (ibid: 34). Relatedly, the UK permitted US rendition flights via British air bases, despite accounts of detainees being transported in sealed, coffin-sized boxes, and SIS officers hearing screams from the hangars (ibid: 32f; European Parliament, 2006).

Over the years, however, legal proceedings and media leaks partially revealed the UK-US collusion to the public, so that in 2010, under the pretext of easing the MI6's workload, David Cameron eventually tasked Peter Gibson with an inquiry into the allegations and introduced a new interrogation guidance[42] (Clarke, 2012). In light of strong criticism, and pending court cases, however, the members of the Gibson Inquiry saw themselves forced to conclude their work prematurely in 2012, publishing a final report and passing on the preliminary results to the Intelligence and Security Committee of Parliament (ISC) (ISC, 2018a: 8). Cameron's 'Consolidated Guidance', in turn, constituted the first publicly available outline of British

interrogation standards for foreigners abroad (HM Government, 2010b); yet, despite this increase in accountability the policy still drew a lot of criticism given its omission of rendition processes, its reliance on assurances, and the considerable discretionary power granted to some of the Ministers (Equality and Human Rights Commission, 2019: 2; Liberty, 2018).

In 2014, the ISC eventually re-opened Gibson's prematurely ended investigation; a step that was originally meant to warrant accountability, but which ultimately bolstered the government's narrative of denial and public pressure mitigation. In light of the "Consolidated Guidance's" limitations, some public actors had attempted to pressure the political leadership into policy-refinement, but the government successfully thwarted such efforts by denying any wrongdoings and by pointing towards pending investigations into any torture-related allegations (Blakely & Raphael, 2020). Similarly, neither leaks of old policies, nor charges of war crimes at the International Criminal Court raised sufficient political or public interest to elicit demands for policy change.[43] The 2015 Report of the Intelligence Services Commissioner Sir Mark Waller posed the only exception, as the critical feedback initiated consultation rounds with the intelligence agencies and Scotland Yard (Cobain, 2018), which, however, were ultimately delayed pending the conclusion of the parliament's ISC report, (ibid.). Finally, also a change of government failed to trigger any type of policy reform, as Theresa May established already early on her opposing stance towards "those activist, left-wing human rights lawyers" (May, 2016), manifesting thus the little role that the torture allegations and corresponding calls for reforms played in her political agenda at that time. In sum, motions within all three streams eventually transpired, yet, none proved strong enough to stipulate the respective streams' ripening.

The Opening of the Problem Window and the ISC Report

The 2018 publication of the Intelligence and Security Committee report, however, marked an important turning point in the reform process of the "Consolidated Guidance" and the "Principles'" introduction as it ultimately triggered the maturing of the politics and problem streams. For four years, the ISC had been investigating torture allegations and issue areas highlighted by the preceding Gibson Inquiry (ISC, 2018a: 9), before then publishing two reports, which on one hand covered the detainee mistreatment between 2001 and 2010, and on the other hand "current issues" (see ISC, 2018a; ISC, 2018b). Even though the reports came with the caveat of little governmental cooperation - 19 of the 23 requested officers had been barred from the witness list (ISC, 2018a: 10) – they confirmed the UK's collusion with the CIA

(ISC, 2018a: 3) and criticized Cameron's 'Consolidated Guidance' as not comprehensive enough (ISC, 2018b: 31-37).

In the end, both documents provided the government with critical and negative feedback, especially as the first report unambiguously states that "it is undeniable that the UK Agencies at Head Office level were aware of reports that some detainees held by the US had been mistreated" (ISC, 2018a: 4). Although the report noted insufficient proof that British intelligence officers were directly involved in torture, it found that detainees had been verbally threatened with mistreatment, and that UK personnel had personally witnessed harmful interrogation techniques inflicted by other countries' agencies (ibid: 3). The ISC report concluded that the deployed staff's lack of experience, and guidance lead to inexcusable actions, which violated international law (ISC, 2018a).

The second report focused on the shortcomings of Cameron's 'Consolidated Guidance', and on whether prevailing orders were formulated clear enough to prevent a repetition of the incidents from the early 2000s (ISC, 2018b). In this context, the ISC voiced major concerns regarding the scope of the 'Consolidated Guidance', as the absence of mandatory policy review processes could result in established rules remaining unenforced (p.2). Similarly, the ISC criticized the high level of ministerial discretion for being a source of subjectivity, while disapproving of the policy's omission of rendition processes. In conclusion, the ISC provided eleven propositions for change and a further 185 pages of closer elaboration on the topic (p. 3).

Despite media coverage and NGO campaigns, the public remained indifferent if not opposed to the ISC publication.[44] Various media outlets covered the findings of the parliamentary investigation, but the lack of graphic evidence as well as the UK's complicit rather than direct involvement in torture hampered any shaming efforts.[45] Similarly, the 2017 scandal surrounding the accusation of fraud against the Iraq Historic Allegations Team had rapidly eroded public trust in comparable inquiries, and instead strengthened the public's loyalty to the British military and intelligence services.[46] This indifference if not resistance to the report combined with the fact of 29% of Brits believing that "torture is sometimes necessary and acceptable to gain information that may protect the public" (Amnesty International, 2014: 5) reduced significantly the likelihood of May's administration suffering electoral losses because of the ISC reports.

Irrespective of the public mood, however, the ISC's negative feedback still caught the government's attention as it reiterated the importance of updating Cameron's 'Consolidated Guidance' in order to prevent the UK's international reputation from being damaged to an

extent where it undermines the state's credibility and threatens national security (ISC, 2018a). One concern pertained to the UK losing its reputation as a 'beacon of human rights',[47] and thus, its legitimacy to act as a mediator or leading voice in international organizations or conflict settings.[48] Furthermore, stakeholders within the government feared that the report's findings might threaten the UK's reputation as a reliable, and human rights respecting partner, which in turn, could drastically endanger any future intelligence cooperation with other countries and weaken the country's negotiation power in future alliance-building processes.[49] Lastly, senior politicians worried that the continued use of a publicly criticized guidance might undermine the military's and intelligence services' trust in the government, thereby undermining the personnel's motivation for defending the country and its international partners.[50]

Yet, it was not only the government itself which grew worried about the ISC reports and their potential implications for the UK on the international stage, but also various Members of Parliament started urging the May Administration for an additional judge-led inquiry and a subsequent policy-reform:

> Furthermore, as we work to ensure that future policy is fit for purpose, it is vital to learn the lessons of the past. The Government's continued failure to deliver an independent inquiry undermines the UK's global reputation as a champion of human rights and the rule of law […]. It is the right thing to do, legally, morally and for British leadership in the world.   (All-Party Parliamentary Group on Extraordinary Rendition, 2019: 2)

In sum, the British government was not under immediate public pressure nor faced immediate threats of electoral losses after the ISC reports' publication. Instead, concerns regarding the UK's reputation, credibility, and negotiation power on the international stage pushed the Consolidated Guidance and its pending reform to the top of the government's problem agenda, signaling thus the ripening of the politics and problem stream.

Stream Alignment and the IPCO Consultations

The same day of the ISC's reports' publication, Theresa May tasked the Investigatory Powers Commissioner's Office to draft an updated policy alternative of Cameron's 'Consolidated Guidance' on interrogations of foreigners abroad (IPCO, 2018a: 3). As a result, IPCO initiated a public consultation process inviting representatives of civil society to submit their policy proposals. The office sent out direct invitations to "key stakeholders", while also using the office's website and Twitter account to openly extend the invitation for submission to "those with an interest in the area" (IPCO, 2018a: 8). The respondents were given three months' time to answer eleven pre-determined questions ranging from their opinion regarding the

Consolidated Guidance to the submission of specific policy proposals. Afterwards, IPCO organized an invitation-only Chatham House event in December 2018 for various NGOs, academics, and government representatives to further discuss the matter from different viewpoints (IPCO, 2018b). Concurrent to these efforts of extracting already well-developed policy alternatives from the policy community, the government likewise ensured that its operational interest was taken into account, and thus, drew the baseline for IPCO's final policy recommendation. Accordingly, the government met repeatedly with IPCO, and likewise responded to the survey, delineating clearly which changes it would consider feasible and which ones not Annex (HM Government, 2018: 5&6).

In the end, IPCO published responses from eight NGOs, the Universities of Sheffield and Westminster, and the All-Party Parliamentary Group on Extraordinary Rendition (IPCO, 2021a). The respondents demanded an absolute prohibition on actions, which carry real risks of torture, renditions as well as cruel, inhumane, or degrading treatment (CIDT), while also demanding the application of the Guidance to scenarios where information sharing with allies would ensue (ibid.). Regarding the Guidance's vague language and the many loopholes the policy provides, all respondents agreed that Cameron's 'Consolidated Guidance' did not provide sufficient direction for (junior) personnel in conflict theaters (ibid). The most prominent points raised were the need for a clear definition of torture and CIDT, an explicit outline of the Minister's competences, as well as detailed and comprehensive risk-assessment procedure applicable to any interrogation scenario (ibid., see responses to the consultation).

In contrast, the meetings between IPCO, the intelligence agencies, and the government representatives remained mostly confidential. Nevertheless, the government answered to IPCO's questionnaire defending the Guidance's scope and precision, but signaled openness to renaming the new policy and including rendition as a form of CIDT into the Annex (HM Government, 2018: 5&6). Regarding the extension of the 'Consolidated Guidance' to the cooperation with non-state actors, the government cited the MI5 claiming that such an alteration was not needed, and instead ensured that "as a matter of policy, the security and intelligence agencies apply the Consolidated Guidance when they know or believe that a detention will take place" (ibid. 5). Similar notions of wanting to retain a high level of flexibility were likewise manifested in the statement that "the government does not consider that the Consolidated Guidance would benefit from being set out on a statutory footing" (ibid: 6).

Policy Adoption

After reconciling the various inputs, IPCO submitted a proposal to Downing Street 10, which the Prime Minister accepted in full (May, 2019). Given May's executive power and the "Principles" non-statutory nature, no further negotiation rounds nor parliamentary voting procedures were needed so that the new policy was adopted as soon as it was amended to May's decision-agenda. As a result, several policy-changes ensued: IPCO was tasked with oversight responsibilities for reports of non-compliance (ibid: 8), while the application of the 'Consolidated Guidance' was expanded to additional UK organizations working in the field of counterterrorism and to international joint units under the lead of UK bodies (ibid: 4&5). Lastly, two of the most salient changes pertain to the inclusion of rendition procedures to the document, and the modification from "serious risk" to "real risk"; both two key demands submitted by civil society actors.

Despite these advancements, however, the Principles do not offer a panacea against torture as they include key objections made by the government, and remain relatively flexible. For example, the cooperation with non-state actors is very vaguely mentioned; in occasions of collaboration the policy "should apply insofar as possible" (ibid: 5). Likewise, other parts of the "Principles" remain vague; Reports of non-compliance should be submitted "as soon as reasonably practicable" (ibid:8), and in cases of time-sensitive conditions "all personnel should continue to observe this guidance insofar as practical" (ibid:7). The government retained a significant amount of leeway; in cases of a "real risk" of torture, CIDT, or rendition, ministers should be consulted, but authorizations were not restricted despite such concerns (ibid:6). Nonetheless, despite all these shortcomings, the Principles still constitute an important advancement compared to the Consolidated Guidance, especially given its inclusion of rendition and the establishment of further oversight responsibilities.

5. Discussion

In light of the policy-making process behind the "Principles", the following section explores and discusses the incongruities between the empirical findings and current MSF literature. In doing so, the article exhibits the limitations of recent theoretical assumptions regarding *consequential coupling* procedures and subsequently refines Herweg et al.'s (2015) hypothesis (H1) to also account for international influences on the opening of problem windows. In a next step, the discussion turns towards Brexit and studies the effects the event has had on the

development of the Principles, before then further exploring May's active involvement in creating a platform for policy-selling efforts.

When looking at the ISC reports' impact on May's policy prioritization, the adherences to the *consequential coupling* logic becomes apparent, even though contrary to Herweg et al's assumption the problem window did not open because of electoral pressures, but rather because of concerns regarding the UK's reputation on the international stage. The ISC reports clearly stated that without policy-change the UK's reputation and credibility would long lastingly be damaged and thwart the state's influence on global governance structures (ISC 2018a&2018b). Now, following Herweg et al.'s (2015: 437) general logic of politicians being inherently interested in securing their own political position and influence, who is to say that such interests are only limited to the domestic realm? Whether it is having a direct influence on how international law is determined or having the capacity to promote and enforce it abroad, politicians both as individuals and representatives of the state have a compelling interest in maintaining their country's access to the international negotiation table. If the country as a whole is being sidelined by other state actors in diplomacy and cooperation processes then this decrease in state power will directly translate into a decrease of the leading politicians' individual power – without access to the negotiation table, the politician's voice will most likely remain unheard.

Building on these inferences, the article proposes a refined hypothesis for the opening of problem windows, which can capture how domestic as well as international issue perceptions can affect the politicians' basic preference of power retention and thus motivate them to adapt their problem agenda accordingly:

H2: The more a condition puts the policymakers' influence on negotiations and general decision-making processes at risk, the more likely it is to open a policy window in the problem stream.

In sum, this hypothesis argues that an issue passes the threshold of opening a problem window when one or more politicians fear to forfeit their say in negotiating generally binding decisions – regardless whether these decisions are being made on the domestic or the international stage. Hence, if applied to the domestic realm, this risk ensues if the politicians' re-election and thus their political position is endangered (see Herweg et al., 2015); in the international realm, however, the risk of losing their say emerges if their personal reputation or the reputation of the state as whole is damaged. Such a loss of credibility could restrict the politicians' access to

international negotiation rounds or, for instance, impede their ability to hold other states accountable.

In this context, the making of the Principles also reveals another phenomenon as during the drafting process, the public mood and the international requirements did seem to oppose each other or at least indicated differing preferences; a conundrum which was mitigated by the side effects of another, at that time prevailing focusing event: Brexit. The UK leaving the European Union had two major effects on British policy-making processes; on one hand the abrupt termination of the almost 50-year long alliance with its European partners made the UK ever more dependent on its international reputation as new partnerships had to be negotiated, while simultaneously trying to limit the diplomatic damages with the European neighbors. On the other hand, Brexit took out all oxygen of political debates, keeping the public distracted from other events on the political stage and thus enabling the government to tacitly pass policy reforms, which were considered as important, but potentially unfavorably viewed by the electorate. Hence, even though Brexit per se was not directly related to the topic of extraterritorial interrogations, it still increased the relevance of impending reputational damages, while providing the decision-makers with sufficient leeway to pursue contentious policy reforms.

In a similar fashion, May's pro-active commissioning of IPCO and the subsequent creation of a platform for policy-selling efforts reflects a part of consequential coupling that is sometimes featured in MSF literature (see Zohlnhöfer, 2016: 97), but not often discussed. One of the basic principles of MSF alludes to the fact that due to the principles of ambiguity and bounded rationality decision-makers have to rely on consultants to obtain feasible policy options when a problem appears (Jones, 2003: 396). In this context, however, most scholars, ascribe a rather passive role to the decision-makers, while the policy entrepreneurs are being portrayed as taking advantage of the politicians' dependency by strategically selling their proposals via framing, moral persuasion, or other approaches (Ackrill & Kay, 2011; Goyal et al., 2020). Yet, in the case of the Principles, May took a pro-active stance when deliberately 'outsourcing' the policy drafting process to IPCO – not to bury the problems in a lengthy bureaucratic process, but to actually facilitate a government-friendly policy reform. Hence, instead of remaining dependent on the policy entrepreneurs' pre-drafted proposals, May rather appointed a third party, which officially is listed as independent, and yet often described as closely attached to the government[51]. Consequently, by getting actively any yet indirectly involved in the solution-finding process, May avoided any potential electoral pushbacks by diminishing her role in the

procedures and by shifting the drafting responsibility towards IPCO; a political move which, given the commissions expertise, also increased the new policy's legitimacy. More importantly, however, by assigning IPCO, May could pre-empt the appointment of a less government-friendly commission or policy entrepreneur, which ultimately enabled her to address the international predicaments, while safeguarding the consideration of the government's operational interests in the new policy draft (IPCO, 2021a).

6. Conclusion

This article has refined Herweg et al.'s (2015) hypothesis on the opening of problem windows by exploring the making of the British Principles, which did ensue due to *consequential coupling* procedures, but in absence of electoral pressures. In doing so, the research demonstrated that the parliament's negative feedback did not provoke discontent among the British public, but rather raised significant concerns regarding the state's reputation and credibility on the international stage, opening thus the corresponding problem window. Being concerned about losing further political power and access to international negotiation tables, Teresa May eventually tasked IPCO with policy finding procedures, securing thus not only the expertise of respective policy entrepreneurs, but also a drafting process within the framework of the government's operational interests. As this development did not fit into the hypothesized casual mechanisms outlined by MSF, the study engaged in theory-building process tracing to present a further refined hypothesis on when a condition passes the threshold of a problem window; an adjustment particularly important when wanting to apply MSF to foreign policy-making processes: The more a condition puts the policymakers' influence on negotiations and general decision-making processes at risk, the more likely it is to open a policy window in the problem stream (H2).

Given the article's reliance on a single-case study, it provides more detailed insights, though until further cross-case studies are done, the generalizability of the findings may be limited. Nevertheless, it can be assumed that the events described in the hypothesis can also occur in other policy fields, as the risk of international reputational-damages is not necessarily unique to the area of human rights and counterterrorism. In light of these considerations, future avenues for research comprise of both theoretical as well as empirical possibilities. Regarding the latter, comparative studies focusing on the emergence of other countries' extraterritorial human rights safeguards could provide further information about respective decision-making processes as

well as their causes, conditions, developments, and outcomes. To that end, a selection of most different cases, including non-western states or non-democratic countries could shed further light on a potential common denominator, which elicits states to recognize extraterritorial human rights obligations. A mix of typical and deviant cases, in turn, could offer details about possible thresholds of recognition. Conversely, theory-focused studies could test the hypothesis in a larger number of cases, engaging thus in large-scale theory-testing via a qualitative comparative analysis, or an outward snowballing-strategy.

7. Bibliography

Abbott, K. W., Keohane, R. O., Moravcsik, A., Slaughter, A. M. & Snidal, D. (2000) 'The concept of legalization', *International Organization*, 54(3): 401-419.

Ackrill, R., & Kay, A. (2011) 'Multiple streams in EU policy-making: the case of the 2005 sugar reform', *Journal of European Public Policy*, 18(1): 72-89.

All-Party Parliamentary Group on Extraordinary Rendition (2019) *Letter to the Prime Minister 28 June 2019,* https://www.extraordinaryrendition.org/documents/appg-letters/category/21-other-correspondence.html [Retrieved October 05, 2021].

Amnesty International (2014) *Global Survey. Attitudes to torture*, https://globescan.com/wp-content/uploads/2017/07/Amnesty-GlobeScan-Global-Survey-Attitudes-to-Torture-2014.pdf [Retrieved March 30, 2021].

Beach, D., & Pedersen, R.B. (2019) *Process-tracing methods: Foundations and guidelines*. Ann Arbor: University of Michigan Press.

Béland, D. (2016) 'Kingdon Reconsidered: Ideas, interests and institutions in comparative policy analysis', *Journal of Comparative Policy Analysis: Research and Practice*, 18(3): 228-242.

Blakeley, R., & Raphael, S. (2020) 'Accountability, denial and the future-proofing of British torture', *International Affairs*, 96(3): 691-709.

Boscarino, J.E. (2009) 'Surfing for problems: Advocacy group strategy in US forestry policy', *Policy Studies Journal*, *37*(3): 415-434.

Brunner, S. (2008) 'Understanding policy change: Multiple streams and emissions trading in Germany', *Global Environmental Change*, 18(3): 501-507.

Cairney, P. & Zahariadis, N. (2016) 'Multiple streams approach: a flexible metaphor presents an opportunity to operationalize agenda setting processes', in N. Zahariadis and M. Buckman (eds.), *Handbook of Public Policy Agenda Setting*. Edward Elgar Publishing.pp.87-104.

Capoccia, G. & Kelemen, R. D. (2007) 'The study of critical junctures: Theory, narrative, and counterfactuals in historical institutionalism', *World Politics*, 59(3): 341-369.

Clarke, K. (2012) *Detainee Inquiry*, https://www.parallelparliament.co.uk/debate/Commons/2012-01-18/debates/12011848000003/DetaineeInquiry [Retrieved March 30, 2021].

Cobain, I. (2018) *Alarm as government rewrites UK 'torture guidance' in secret*, https://www.theguardian.com/law/2018/apr/01/alarm-as-government-rewrites-uk-torture-guidance-in-secret [Retrieved March 30, 2021].

Copeland, P., & James, S. (2014) 'Policy windows, ambiguity and Commission entrepreneurship: explaining the relaunch of the European Union's economic reform agenda', *Journal of European Public Policy*, 21(1): 1-19.

European Parliament (2006) *Alleged secret detentions and unlawful inter-state transfers of detainees involving Council of Europe member states*, https://assembly.coe.int/CommitteeDocs/2006/20060606_Ejdoc162006PartII-FINAL.pdf [Retrieved March 30, 2021].

European Parliament (2007) *Report on the alleged use of European countries by the CIA for the transportation and illegal detention of prisoners (2006/2200(INI))*,

https://www.europarl.europa.eu/comparl/tempcom/tdip/final_report_en.pdf [Retrieved March 30, 2021].

Equality and Human Rights Commission (2019) *Torture in the UK: Updated Report,* https://www.equalityhumanrights.com/sites/default/files/convention-against-torture-in-the-uk-executive-summary-may-2019.pdf [Retrieved March 30, 2021].

Farley, J., Baker, D., Batker, D., Koliba, C., Matteson, R., Mills, R. and Pittman, J. (2007) 'Opening the policy window for ecological economics: Katrina as a focusing event', *Ecological Economics*, 63(2-3): 344-354.

Fariss, C. J. (2014) 'Respect for human rights has improved over time: Modeling the changing standard of accountability', *American Political Science Review,* 108(2): 297-318.

Geddes, B. (2003) *Paradigms and sand castles: Theory building and research design in comparative politics*, Ann Arbor: University of Michigan Press.

Gerring, J. (2004) 'What Is a Case Study and What Is It Good for?' *The American Political Science Review*, 98(2): 341-354.

Goyal, N., Howlett, M. and Chindarkar, N., (2020), 'Who coupled which stream (s)? Policy entrepreneurship and innovation in the energy–water nexus in Gujarat, India', *Public Administration and Development*, 40(1): 49-64.

Herweg, N., Huß, C. and Zohlnhöfer, R. (2015) 'Straightening the three streams: Theorising extensions of the multiple streams framework', *European Journal of Political Research*, 54(3): 435-449.

Herweg, N., Zahariadis, N., & Zohlnhöfer, R. (2018) 'The multiple streams framework: Foundations, refinements, and empirical applications', in C.M. Weible and P.A. Sabatier (eds.), *Theories of the Policy Process*. New York: Routledge. Pp. 17-53.

HM Government (2010a) *Consolidated Guidance to Intelligence Officers and Service Personnel on the Detention and Interviewing of Detainees Overseas, and on the Passing and Receipt of Intelligence Relating to Detainees,*https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attach ment_data/file/62632/Consolidated_Guidance_November_2011.pdf [Retrieved March 30, 2021].

HM Government (2010b) *Note of Additional Information from the Secretary of State for Foreign and Commonwealth Affairs, the Home Secretary, and Defence Secretary*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/ file/62633/additional-information-fcohods_0.pdf [Retrieved March 30, 2021].

HM Government (2018) *IPCO Consultation on the Consolidated Guidance. HM Government Response,* https://www.ipco.org.uk/docs/HMG.pdf [Retrieved March 30, 2021].

HM Government (2019) *The Principles relating to the detention and interviewing of detainees overseas and the passing and receipt of intelligence relating to detainees*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/ file/818306/20190718_The_Principles_relating_to_the_detention_and_interviewing_of_detai nees_overseas.pdf [Retrieved March 28, 2021].

IPCO (2018a) *Consultation on the Consolidated Guidance*, https://ipco-wpmedia-prod-s3.s3.eu-west-2.amazonaws.com/IPCO-Consultation-on-the-Consolidated-Guidance1.pdf [Retrieved March 30, 2021].

IPCO (2018b) *Summary Chatham House Discussion,* https://ipco-wpmedia-prod-s3.s3.eu-west-2.amazonaws.com/IPCO-Statement-on-the-Chatham-House-discussions-12-Dec-2018.pdf [Retrieved March 30, 2021].

IPCO (2021a) *Consultations*, https://www.ipco.org.uk/publications/consultations/consultations-documents/page/2/ [Retrieved March 30, 2021].

IPCO (2021b) *Who we are*, https://www.ipco.org.uk/who-we-are/ [Retrieved March 30, 2021].

Intelligence and Security Committee of Parliament (2018a) *Detainee Mistreatment and Rendition_ 2001-2010*, https://fas.org/irp/world/uk/isc-detainee.pdf [Retrieved March 30, 2021].

Intelligence and Security Committee of Parliament (2018b) *Detainee Mistreatment and Rendition_ Current Issues*, https://fas.org/irp/world/uk/isc-detainee2.pdf [Retrieved March 30, 2021].

Jones, M.D. (2003) 'Bounded rationality and political science: Lessons from public administration and public policy', *Journal of Public Administration Research and Theory*, 13(4): 395-412.

Kingdon, J. W. (1984) *Agendas, alternatives, and public policies*, Boston: Little, Brown.

Knaggard,A. (2015) 'The Multiple Streams Framework and the Problem Broker', *European Journal of Political Research*, 54(3): 450-465.

Liberty. (2018) *Liberty's response to the Investigatory Powers Commissioner's consultation on Consolidated Guidance*, https://www.ipco.org.uk/docs/Liberty.pdf [Retrieved March 30, 2021].

May, T. (2016) *Keynote speech at Tory conference in full*, https://www.independent.co.uk/news/uk/politics/theresa-may-speech-tory-conference-2016-full-transcript-a7346171.html [Retrieved May 5, 2021].

May, T. (2018) *Statement*, https://questions-statements.parliament.uk/written-statements/detail/2018-06-28/HCWS808 [Retrieved March 30, 2021].

May, T. (2019) *Review of Consolidate Guidance. Statement made on 18 July 2019,* https://questions-statements.parliament.uk/written-statements/detail/2019-07-18/HCWS1738 [Retrieved March 30, 2021].

Miller, T., Birch, M., Mauthner, M. and Jessop, J. (2012) *Ethics in Qualitative Research.* London: Sage Publication.

Rohlfing, I. (2014) 'Comparative hypothesis testing via process tracing', *Sociological Methods & Research*, 43(4): 606-642.

Siggelkow, N. (2007) 'Persuasion with case studies', *The Academy of Management Journal*, 50(1): 20-24.

Sikkink, K. (2013) 'The United States and torture: Does the spiral model work?', *Cambridge Studies in International Relations,* 126(1): 145-163.

Travis, R. and Zahariadis, N. (2002) 'A Multiple Streams Model of U.S. Foreign Aid Policy', *Policy Studies Journal*, 30(4): 495-514.

UK Parliament (2019) *Review of Consolidated Guidance. Statement made on 18 July 2019*, https://questions-statements.parliament.uk/written-statements/detail/2019-07-18/HCWS1738 [Retrieved March 30, 2021].

UK Parliament (2021) *UK war crimes in Iraq: The ICC Prosecutor's report,* https://commonslibrary.parliament.uk/uk-war-crimes-in-iraq-the-icc-prosecutors-report/ [Retrieved March 30, 2021].

van der Heijden, J., Kuhlmann, J., Lindquist, E., Wellstead, A. (2019) 'Have policy process scholars embraced causal mechanisms? A review of five popular frameworks', *Public Policy and Administration*, 36(2): 163-186.

Zahariadis, N. (2003) *Ambiguity and choice in public policy: Political decision making in modern democracies*. Washington D.C.: Georgetown University Press.

Zahariadis, N. (2016) 'Delphic oracles: Ambiguity, institutions, and multiple streams', *Policy Sciences*, 49: 3-12.

Zhu, X. (2008) 'Strategy of Chinese policy entrepreneurs in the third sector: Challenges of technical infeasibility', *Policy Sciences,* 41(4): 315-334.

Zohlnhöfer, R., Herweg, N. and Huß, C., (2016) 'Bringing formal political institutions into the multiple streams framework: An analytical proposal for comparative policy analysis', *Journal of Comparative Policy Analysis: Research and Practice*, 18(3), pp.243-256.

Zohlnhöfer, R. (2016) 'Putting together the pieces of the puzzle: Explaining German labor market reforms with a modified multiple-streams approach', *Policy Studies Journal*, 44(1): 83-107.

---

[39] "Principles relating to the detention and interviewing of detainees overseas and the passing and receipt of intelligence relating to detainees" (henceforth "Principles").

[40] Independent oversight body of the British intelligence agencies (IPCO, 2021b).

[41] In the context of MSF, negative feedback is provided by internal or external actors to the government, and refers to among others, discrepancies between the implementation and the original policy intent, a failure in meeting the stated goal, unexpectedly high costs, or unanticipated consequences (Zohlnhöfer et al., 2016: 255; Kingdon, 1984: 102f).

[42] Consolidated Guidance to intelligence officers and service personnel on the detention and interviewing of detainees overseas, and on the passing and receipt of intelligence relating to detainees; henceforth referred to as "Consolidated Guidance".

[43] Interview 3 with former General Counsel for the Equality and Human Rights Commission, London 17/02/2020.

[44] Interview 4 with former Reader in Foreign Policy and International Relations University of Birmingham, Birmingham 05/03/2020.

[45] Interview 1 with Senior Member UK National Preventive Mechanism, London 11/02/2020.

[46] Interview 1.

[47] Interview 5 with a senior scholar of the University of Westminster, London 14/02/2020.

[48] Interview 6 with member of the United Nations Subcommittee for the Prevention of Torture, London 25/02/2020.

[49] Interview 7 with former UK Security and Intelligence Coordinator for the Prime Minister, London 04/03/2020.

[50] Interview 8 with the former Attorney General for England and Wales, London 27/04/2020.

[51] Interview 5 with a senior scholar of the University of Westminster, London 14/02/2020.

Annex:

1. *List of Interviewees*

Interviews USA

19 March 2019: Senior Fellow, Brookings Institution, Washington D.C.

19 March 2019: Associate Professor, Georgetown University, Washington D.C.

19 March 2019: Former CIA Agent, Washington D.C.

20 March 2019: Washington Director, Center for Victims of Torture, Washington D.C.

20 March 2019: Retired Admiral, US Navy, Washington D.C.

25 March 2019: Deputy Washington Director, Human Rights Watch, Washington D.C.

25 March 2019: Retired Brigadier General, US Army, Washington D.C.

28 March 2019: Former President and CEO, Human Rights First, Washington D.C.

28 March 2019: Retired Rear Admiral, US Navy, Washington D.C.

28 March 2019: Retired US Army officer, Skype.

29 March 2019: Retired Lieutenant General, US Army, Washington D.C.

2 April 2019: Former Head of the Strategic Interrogation Program, US Air Force, Skype.

4 April 2019: Former General Counsel of the US Navy, Washington D.C.

8 April 2019: Former Member of the President's Intelligence Advisory Board and Executive Director of the 9/11 Commission, Skype.

9 April 2019: Former US Special Representative for Afghanistan and Pakistan, Washington D.C.

10 April 2019: Former Senior Associate Counsel to the President and Legal Adviser to the National Security Council, Washington D.C.

10 April 2019: Former CIA Agent and Deputy National Intelligence Officer for Transnational Threats on the National Intelligence Council, Skype.

12 April 2019: Spokesperson Amnesty International, Washington D.C.

25 April 2019: Former Director, Center for Victims of Torture, Skype.

3 May 2019: Former Director of the Criminal Investigative Task Force, Department of Defense, Skype.

1 July 2019: Former Assistant Secretary of State, Bureau of Democracy, Human Rights and Labor, Skype.

Interviews Germany

03 September 2019: Spokesperson Human Rights Watch Germany, Skype.

09 September 2019: Former UN Special Rapporteur on Freedom of Religion or Belief, Berlin.

10 September 2019: Former Federal Commissioner for Human Rights Policy and Humanitarian Assistance, Berlin.

10 September 2019: Former Chair Bundestag Committee of Human Rights and Humanitarian Aid, Berlin.

12 September 2019: Former Director German Institute for Human Rights, Berlin.

13 September 2019: Former Member BND-Untersuchungsausschuss, Berlin.

19 September 2019: Former Deputy Chair UN Committee on Civil and Political Rights, Berlin.

Interviews UK

11 February 2020: Senior Member UK National Preventive Mechanism, London.

12 February 2020: Member of Parliament, Conservative Party, London.

14 February 2020: Spokesperson Reprieve, Skype.

17 February 2020: Former General Counsel for the Equality and Human Rights Commission, London.

19 February 2020: Member of Parliament, Labour Party, London.

20 February 2020: Member of the Privy Council and Former Member Gibson Inquiry.

20 February 2020: Senior Scholar for International Law at University of Birmingham, Birmingham.

25 February 2020: Former Member of the United Nations Subcommittee for the Prevention of Torture, London.

26 February 2020: Former Chair of the UK's National Preventive Mechanism, London.

27 February 2020: Former Member Iraq Historic Allegations Team, Skype.

04 March 2020: Former UK Security and Intelligence Coordinator for the Prime Minister, London.

05 March 2020: Former Reader in Foreign Policy and International Relations University of Birmingham, Birmingham.

27 April 2020: Former Attorney General for England and Wales, Skype.