

Scorecard Diplomacy and Reputation

It was really a reputational issue. Kazakhstan wanted very much to be seen as a modern country, or it has aspirations for that. It wanted to be accepted.

– *Larry Napper, Ambassador to Kazakhstan*¹

Trafficking in persons was not a big priority for Kazakhstan’s government in the late 1990s. In 1999 the government’s National Commission on Women’s and Family Issues even declined to include trafficking in its list of priorities. However, after the 2001 TIP Report labeled Kazakhstan as Tier 3, interactions with the embassy spiked as high-level officials became concerned about this bad rating.

The use of public grades is a core element of scorecard diplomacy. This book argues that these grades are central in making countries receptive to broader diplomatic engagement. In a world more accustomed to “muscle diplomacy” this is odd. Why should meager grades make a difference, especially on issues that are not high profile? Why should countries care about a grade, especially if it is good enough to avoid consequences such as aid loss? And perhaps most puzzling: why should countries care what grades *other* countries get?

The answers to these questions lie in the concept of reputation. This chapter delineates the role of reputation in states’ behavior and explains how the features of scorecard diplomacy raise states’ concerns about their reputations. It argues that states care about their reputation not simply in terms of the credibility of their promises and threats, but more broadly in terms of how others perceive their performance. States want social recognition and governments care about their reputation vis-à-vis

¹ Napper interview.

their citizens and the global community because this directly affects their standing and legitimacy to govern.²

This desire means that others can seek to influence states by influencing their reputation. As others have noted, power “is a social process of constituting what actors are as social beings, that is, their social identities and capacities” to determine their fate.³ While this has always been true, the digital age has made it cheaper to generate and disseminate criticism and accountability systems have become more transparent.⁴ Thus, the ability to influence the reputation of states is an increasingly valuable tool.

Reputational concerns are only one of many factors that influence states. Sometimes it carries more weight than others. Reputational concerns won't compel North Korea's Kim Jong-un to resign or stop human rights repression in China, but it curbed the ability of the US to use torture and has tempered countries' trade in weapons.⁵ Furthermore, even if reputational concerns may not *constrain* China's ability to repress human rights, it still imposes inconvenient *costs* that China must factor into its interstate relations.

This chapter has three parts. It first provides a broad definition of reputation and argues why it is valuable to states. It then creates a simple model to explain when states are more likely to worry about their reputation and act accordingly. Finally, it explains how scorecard diplomacy stimulates countries concern with their reputation.

* * *

THE CONCEPT OF REPUTATION IN INTERNATIONAL RELATIONS

Some scholars argue that reputation matters “most in trade and security and least in environmental regulation and human rights”⁶ because misbehaving governments know that while important allies or trade partners prefer to tolerate violations on soft issues like human rights rather than let misconduct spill over into higher priority areas like trade and security.⁷

² Wendt 1999, Ch. 5.

³ Barnett and Duvall 2005, 42.

⁴ Nye 1990, 100, Grant and Keohane 2005.

⁵ Erickson 2015.

⁶ Downs and Jones 2002, 5112.

⁷ Downs and Jones 2002, 597. Simmons discusses, but does not endorse, these views. Simmons 2009, 122.

However, this pessimistic view rests on too narrow a use of the word reputation that equates reputation with “credibility of commitments,” or a reputation for resolve not to back down in the face of opposition.⁸ Following this hard “reputation-as-credibility” logic, a reliable reputation enables states to gain from repeated cooperation with other states.⁹ Reputation thus defined is about the predictability of a state’s actions and it signals a state’s military resolve or reliability as a partner on trade issues and so forth.¹⁰ Indeed, reputation has become so synonymous with credibility of commitment that many use the terms almost interchangeably.¹¹ Even if reputation is defined more broadly as general beliefs about an actors’ behavior, its consistently applied narrowly to topics such as threats, retaliation, reliability as an ally, or treaty compliance, perhaps because the study of reputation came of age during the Cold War and the study of deterrence.¹² Some even call this the “rational dimension of reputation that is chiefly of interest to economists and most political scientists,” as if other dimensions of reputation are irrational or of less interest.¹³ It is unfortunate that the concept of reputation has become shorthand for this narrower meaning.¹⁴ As a result, much of the empirical research on “reputation” has focused on issues that involve iterated cooperation, such as security and economic issues.¹⁵

⁸ Kreps and Wilson 1982, Alt et al. 1988, Jervis 1989, Sharman 2007. On resolve, see Dafoe et al., 375.

⁹ Keohane 1984, 94, Schelling 1980.

¹⁰ See in particular Mercer 1996, 6. Also Mercer 1996, Sartori 2002, Crescenzi 2007, Crescenzi et al. 2007, Walter 2009, Levy and Thompson 2011, Hugh-Jones 2013. Keohane notes that “[t]o a government that values its ability to make future agreements, reputation is a crucial resource; and the most important aspect of an actor’s reputation in world politics is the belief of others that it will keep its future commitments.” Keohane 1984, 116. For a treatment on sanctions see Peterson 2013. For reputation and sovereign debt, see Tomz 2007b. This conceptualization figures in some work on legalization and international law. See Abbott and Snidal 2000, Simmons 2000, 2010. On “Reputational capital,” see Guzman 2002.

¹¹ Tomz 2007b. Even domestic audience costs have been conceptualized as a reputational cost of breaking commitments or being caught bluffing. Fearon 1994, Sartori 2002, Tomz 2007a. Joseph Nye also uses “reputation” to refer to credibility. Nye 2008, 100.

¹² See for example Dafoe et al. 2014.

¹³ Downs and Jones 2002, S96, fn 2.

¹⁴ Dafoe et al. 2014, 375. Even scholars those who disagree with these schools of thought use of the word reputation to refer to credibility. See Johnston 2008, 7.

¹⁵ Mercer 1996. I myself have used a narrow definition of reputation as synonymous with concerns about the benefits of future exchange, a view I now consider too narrow. Kelley 2004b, 428.

A Broader Definition of Reputation

In the spirit of Malthus who advocated that scholars use words according to their common understanding, this book uses the word reputation broadly and more consistently with everyday usage.¹⁶ Rather than being foremost or only about credibility of threats or promises, reputation here refers to the basic Merriam-Webster dictionary definition as the “the common opinion that people have about someone or something,” or “the overall quality or character as seen or judged by people in general.”¹⁷ That is, a state’s reputation reflects how others view it or – where responsibility can easily be allocated – its officials.¹⁸ Reputations are usually based on past behaviors and can be updated with new information that indicates changes in the track record.¹⁹ Reputation in this broad sense is not just about keeping promises; it’s about perceptions of performance, both in terms of processes and outcomes across a range of issues. Thus, reputation is about character more broadly. Such a conceptualization that includes a notion of reputation as image or status has long existed. It is akin to what some now call “social reputation,” or “diffuse reputation or image,” which is the “the package of favorable perceptions and impressions that one believes one creates through status consistent behavior.”²⁰

Reputation is not a fixed property but is in the eye of the beholder. States can have reputations in relation to multiple actors: citizens, national elites, other governments, and the global community. These audiences

¹⁶ Thomas Malthus, writing in 1827, thought it incumbent upon social scientists seeking to use common words to denote concepts to use those words in ways that accord with common usage. Those who narrowly equate reputation with credibility of commitments, etc., violate this basic advice. Thanks to Baldwin 2016, 44–45, for highlighting these insights of Malthus.

¹⁷ See also Sharman 2007, Erickson 2015. Erickson (Ch. 2) has an outstanding discussion of reputation as “image and social status,” or “credible cooperation,” or “credible threat.” See in particular her Table 2.1. For a similar broad definition, although their application is more narrowly to conflict, see also Dafoe et al. 2014.

¹⁸ Indeed, much theory on reputation and image moves easily back and forth between analysis of individual and state level. A large debate exists about whether reputations are attached to states or leaders, but this most likely depends on whether the issue and related decisions are diffuse or can easily be identified with individuals who can be held accountable. For a longer discussion and a survey experiment backing this intuition in the context of war, see Renshon et al. Forthcoming.

¹⁹ Tomz 2007b, 17–20.

²⁰ For use of the term “social reputation,” see Risse and Ropp 2013, Erickson 2015. For the notion of “diffuse reputation or image,” see Johnston 2001, 500, 2008. See also related discussions of status as communal proficiency assessments. For a discussion of the relationship between image and reputation, see Erickson 2015, 25–26.

may assess a state's reputation differently depending on which standards they each use, what they know, or how they weigh the evidence.²¹

Reputation, Behavior, and the Existing Normative Environment

Widely accepted rules or practices that prescribe appropriate or desirable behavior are central to how states and governments assess each other, and how citizens assess their governments.²² One can think of reputation in an idealized form. In Figure 2.1, point "I" represents the internationally defined ideal behavior along some spectrum. For example, this could be the degree of freedom of speech, policies to promote gender equality, or per capita carbon emissions. The ideal point could be embedded in an international agreement or in common practices. It could be about policy behavior, but it could also be about behavioral traits, such as generally abiding by the rule of law, or keeping international commitments. The more the global community agrees on and adheres to the ideal standards, the greater their weight and the more these standards act as prerequisites for international legitimacy.²³

Based on Figure 2.1, a country's international reputation is the way the international community at large assesses the gap between ideal point "I" and the state's actual behavior, "B." The smaller the gap, the better the reputation. Countries have a "good" reputation when they approximate "I." In reality of course, "I" might be more or less well established, and the understanding of "B" may vary with the level and quality of information, but the idea is that the international reputation is the assessment of the gap by international audiences, and this assessment is more favorable the smaller the gap.

Importantly, however, this reputation may or may not be the same as the one the government has at home. If the state subscribes to different ideals than the international community, domestic audiences may assess

²¹ Tallberg and Zürn 2015.

²² For a broader discussion of the role of a constitutive environment, see Wendt 1999. Baldwin links the constructivist notions about a socially constructed reality to earlier ideas by Harold Lasswell and Abraham Kaplan about power being "situational." Baldwin 2016, 153–154.

²³ Complete consensus around norms need not exist for states to gain reputations in a given issue area. Different ideologies or competing norms may dictate different behaviors and different groups may judge states differently depending on which ideologies they subscribe to. For example, some states may have a reputation as abiding by the principles of the free market, while others may have a reputation for managing their economies more closely. In this case reputations are not necessarily measured against a universal right or wrong, but are labels for different – sometimes competing – ideologies.

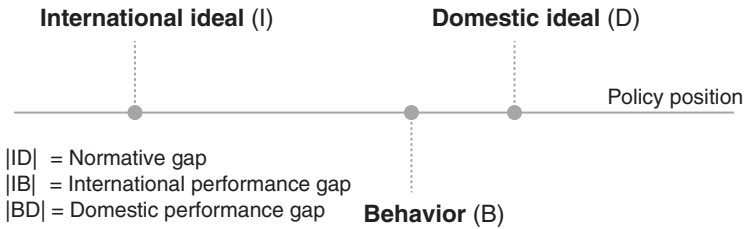


FIGURE 2.1. Performance gaps and the relationship between ideals and practice.

the state's behavior differently. For example, in some countries a widespread acceptance of female circumcision differs from the more widely held international taboos. In Figure 2.1, "D" represents the domestic ideal point. If the international norms conflict with more locally held norms, then "I" and "D" will be far apart and the state may have a different reputation at home than abroad because of the assessment of the gap between the domestic ideal point and behavior differs.

This can place a government in a conundrum. Internationally it is concerned about the gap between its behavior and the international idea, $|IB|$, while domestically it is concerned about $|BD|$, the domestic performance gap. Naturally, "D" could also lie between "I" and "B," in which case the direction of the gap is the same, representing less of a dilemma, but still leading to differential pressure domestically and internationally. Or "D" and "I" could coincide, in which case international criticisms ring true at home. In either case, reputations form relative to existing ideal points and behavior. Because there can only be one behavior, differing ideal points may pull government policies in opposite directions.

Importantly, ideal points need not be fixed. The very act of invoking certain ideals can be an exercise not only in rule enforcement but also in standard setting. That is, those assessing performance are simultaneously shaping the interpretation of what is acceptable performance. In this way, invoking norms is also an act of influencing the very definition of these standards and norms.²⁴ Invoking the norms and calling for their implementation can help institutionalize the norms and standards.²⁵ Indeed, information is powerful because it promotes accountability, but

²⁴ This is akin to the argument that contestation over norms shapes their form. Finnemore 1996. Much of the literature on governance indicators engages this argument. See for example Löwenheim 2008, Merry 2011, Büthe 2012, Davis et al. 2012a, Kelley and Simmons 2015, 2014, Merry et al. 2015.

²⁵ Finnemore and Sikkink 1998.

also because the producers of the information and assessments influence the definition and salience of ideas.²⁶

Multiple Reputations Across Issues

In accordance with this broader definition, countries not only have multiple reputations across different audiences, but also across different issues or traits.²⁷ For example, countries that stimulate positive economic growth, lead humanitarian rescue efforts, or advance international peace talks may earn a good reputation in the view of those who favor such policies. Conversely, countries that monopolize domestic markets, confiscate private property, or slash and burn its virgin forests might earn a poor reputation in view of those opposed to such practices. Generally, huge environmental disasters tarnish a government's reputation as a competent and responsible caretaker. The selling of arms to human rights abusers brands a government as opportunistic in the eyes of most. Conduct in trade and war will brand a government's reputation for credibility and resolve.

While states have multiple reputations, they also accumulate a general reputation based on their performance across many issues ranging from the provision of services and the stewardship of resources to various international matters. As others have noted, "to say that a state has a particular reputation implies that most observers hold the relevant belief about the state."²⁸ In this sense, states may "acquire reputations as law-abiding global citizens."²⁹

Importantly for whether countries are concerned about their reputation in a given issue area, many issues cannot be compartmentalized. If countries repeatedly disregard negative externalities they impose on their neighbors in different areas, for example, other states may worry that poor consideration for neighbors is a general quality.³⁰ In an age when states are considering much broader approaches to "human security," this is increasingly true. Thus, the potential for spillovers remains for most issue areas. This means that a state's overall reputation can be damaged by a poor performance in any issue area that reveals a fundamental

²⁶ Keohane and Nye Jr 1998, 86, Kelley and Simmons 2014.

²⁷ For a discussion stressing this, see Dafoe et al. 2014, 374.

²⁸ Dafoe et al. 2014, 374.

²⁹ Lebovic and Voeten 2006, 885.

³⁰ Crescenzi argues for an experiential model of reputation in which states learn from how other states behave more broadly in the international community, not just from dyadic interactions. Importantly, the way that states assess information depends on the similarity of context. Crescenzi 2007.

disregard for qualities of broader universal values.³¹ This can be true even if the issue area itself is rather narrow. For example, in 2005 when Mugabe razed local slums in “Operation Murambatsvina,” his reputation suffered generally, not only on the issue of local housing, or even human rights more generally, but also on his governance more broadly. He was widely seen as ruthless and power-obsessed and his state as edging ever closer to pariah status, as noted by a condemning report from the United Nations.³² Just like bad news tends to dominate the media, so government failures that attract considerable attention tend to dominate their reputations. States therefore have reason to be concerned about their reputation in any given issue area that threatens to attract public attention.

Furthermore, as opposed to a purely credibility-based conceptualization of reputation, a broader performance and image-based conceptualization sees human rights as central, because it involves moral responsibility and justice, which are the fundamental building blocks of state “character.” Maintaining legitimacy requires states to conform to the international community’s minimal justice requirement.³³ Because of the substantive international support for human rights norms, countries that violate human rights struggle to maintain legitimacy.³⁴ Thus, states may value having a compliant reputation in areas such as human rights that undergird their overall character.

Why Care? The Value of a Good Reputation

A narrow definition of reputation stresses that states value their reputation because they want to gain from cooperation or be able to make credible threats in wartime. In the context of the latter one might allow that some leaders actually value a reputation for unpredictability, ruthlessness,

³¹ See Erickson’s excellent discussion of scandal and domestic politics. Erickson 2015, 31–34.

³² Tibajjuka 2005.

³³ Barkin and Cronin 1994, 113, Buchanan 2003, 282. States are considered responsible when they observe human rights including broad social and economic rights. Franck 1990, Wendt 1999, Lebovic and Voeten 2006. Thus, countries that violate human rights struggle to maintain legitimacy.

³⁴ Hawkins 2002, 31. States cannot maintain their own image as upright if they ignore blatant human rights violations in their own or other countries. Thomas 2001, 51. Indeed, theories of shaming assume that states do care about their human rights reputation and that elites worry about others’ assessment of their performance or their character. Lebovic and Voeten 2006, Franklin 2008, Efrat 2009.

and so forth. That is, some leaders might prefer a bad reputation – they actually wish for their behavior to be assessed as far from the international ideal point.

That said, a broader definition also stresses that governments and citizens value their reputation as part of their state's identity or image, similar to how a citizen may desire to be viewed as law-abiding.³⁵ If states have a good reputation – if others assess their behavior as aligned with international ideals – this gives states and their governments a sense of belonging, facilitates cooperation with other states, and allows them to consider themselves as upright members of the international community.³⁶ A positive reputation can be a policy goal in itself, not merely a means to an end.³⁷ Indeed, at times states may value their reputation so highly that they are willing to forego other immediate or more concrete gains.³⁸

That states and elites care about their reputation in terms of image and broader legitimacy does not imply that states are not instrumental or strategic. States may adopt certain norms and behaviors to reap social benefits and avoid social costs; in this sense, they respond rationally to social incentives.³⁹ A good reputation can confer respect and influence, which enables states to accomplish other goals.⁴⁰ Thus, for states, concern

³⁵ Lutz and Sikkink have argued that “as members of an international or regional society of states [leaders of authoritarian governments have] been ‘socialized’ into caring about what other states think of them.” Lutz and Sikkink 2000, 659. See also Finnemore 1996, Risse et al. 1999, Erickson 2015. In prior work I have argued that citizens may expect their governments to respect international law as part of their state's identity. Kelley 2007.

³⁶ Thus, Franck argues that states seek to become members of a club of states, and others argue that shaming is meaningful because it places states in an out-group, and causes some states to “feel deeply offended, because they want to belong to the ‘civilized community’ of states.” Franck 1990, 38. In international relations, identity or image is often related to “social identity theory” and used to denote the idea of belonging to certain groups, and to the concept of “othering,” as part of a state's own identity construction. Risse and Sikkink 1999, 15.

³⁷ Erickson 2015, 24. Johnston proposes that “actors in world politics value image and status as ends in and of themselves.” Johnston 2008, 75.

³⁸ Johnston notes that often status markers are uncorrelated with material gains, and desires to maximize status need not even be about topping others. Johnston 2008, 83. Others note that “reputation is often crucial for status. The loss or acquisition of certain reputations can lose or gain an actor a particular status.” Dafoe et al. 2014, 375.

³⁹ For a discussion of social incentives, see Erickson 2015, 17–24. Both Hurd and Johnston note that states act rationally within a social context. Hurd 2008, Johnston 2008. Kelman refers to this behavior simply as compliance. Kelman 1958.

⁴⁰ Gilpin 1981, 30–31.

about their image “is always in the national interest.”⁴¹ In other words, a good reputation is just plain useful.

Favorable reputations boost states’ legitimacy, both internationally and domestically. “International legitimacy” is coveted because it is foundational to the state and international society and underscores the right to govern.⁴² Even strong states and their leaders seek international recognition and may assign value to upholding norms that are widely accepted in the international community – a point that some have argued explains why many states joined in sanctions against South African apartheid.⁴³ The US, for example, has worried about anti-Americanism worldwide and has sought to manage its international reputation.⁴⁴

A favorable overall reputation is also useful domestically.⁴⁵ Citizens assess their states both for their procedural legitimacy and for the outcomes they produce. It is procedural reasons that have compelled states to invite international monitors to verify the legitimacy of electoral processes.⁴⁶ If citizens perceive their government as corrupt or irresponsible, they may lower their support for the government.⁴⁷ States and their elites may fear economic repercussions or worry about their own political survival.⁴⁸ Although most violations of international norms won’t topple the government, cumulative misconduct can erode its reputation over time. For example, the revelation of human rights violations by the US military in the Abu Ghraib prison in Iraq brought protests at home and undercut

⁴¹ Baldwin 2016, 171. For the value of national image, see also Jervis 1989. Public criticism can harm a state’s ability to cooperate with other states and can damage its domestic authority to govern. Gilley 2013.

⁴² On the role of legitimacy in international society, see Bull 1977, Hurd 1999, Clark 2005. International legitimacy, too, is a social concept – in the eyes of the beholders – and denotes the extent to which others accept the *authority* of a regime and view it as having the right to govern. Reus-Smit discusses legitimacy as a social concept and argues that it cannot be separated from the concept of power, but is a source of power. Reus-Smit 2007.

⁴³ Klotz 1995.

⁴⁴ Johnston 2008. On anti-Americanism, see Katzenstein and Keohane 2007. On US standing, see American Political Science Association 2009.

⁴⁵ See Lutz and Sikkink 2000, Newton and Norris 2000, Reus-Smit 2007, Hurd 2008, 29–30, Kelley 2012, Risse and Ropp 2013. Wendt notes that legitimacy derives from possessing “identity criteria which define only certain *forms* of state as legitimate.” Wendt 1999, 292. Erickson has a particularly extensive discussion of the importance of national reputation. Erickson 2015, Ch. 2.

⁴⁶ A survey of eight African countries shows that individuals who perceive election conduct as proper are more likely to consider the regime legitimate. Elklit and Reynolds 2002. See also Kelley 2012, Ch. 2.

⁴⁷ For a discussion of procedural and performance legitimacy as it pertains to international organizations, see Tallberg and Zürn 2015.

⁴⁸ Batson 1987.

support for the president. Thus, governments need good reputations to maintain domestic support.

Although some dictators may seek to inspire fear by deviating from international norms, support is important for all regime types. Even authoritarian governments experience pressures to perform and are vulnerable to external criticisms, which is one reason so many of them hold “elections” and invite international election monitors to gain some sheen of legitimacy.⁴⁹ Indeed, because authoritarian states compromise on procedural legitimacy, some governments, such as Singapore and China, rely on a reputation of being able to deliver certain policy outputs to sustain their legitimacy.⁵⁰ For example, Singapore prides itself on its top placement in the Heritage Foundation’s Economic Freedom Index as a form of validation, which is why the government has worked so hard to maintain the top spot for decades and brags about it.⁵¹ Indeed, some research has found that authoritarian states are more susceptible to public exposure of human rights violations than are democracies.⁵² Naturally, authoritarian regimes may compare themselves with different “reference” states, but the principle remains if for some issues they wish to be viewed favorably by a peer group or by domestic constituents.

In addition to the benefits of domestic and international legitimacy, a good reputation also boosts status and standing in the international community. A state’s status refers to its position relative to a comparison group,⁵³ and denotes a more hierarchical structure of state relations. Higher social status and prestige are valuable as means to boost recognition and influence in international society or to attract investment or other benefits.⁵⁴

Although the discussion here has mostly focused on states and their governments, individual leaders and elites also have reasons to worry about their *personal* reputation. Most want to be respected by their citizens and leaders of other states with which they identify.⁵⁵ They want their country to compare well with other countries and they shun “the

⁴⁹ Hawkins 2002, 30, Johnston 2008. On election monitoring, see Hyde 2011, Kelley 2012.

⁵⁰ Zhu 2011.

⁵¹ Interview with staff, by Judith Kelley, August 12, 2014, Washington, DC.

⁵² Hendrix and Wong 2013.

⁵³ Dafoe et al. 2014.

⁵⁴ Gilpin 1981, 30–31, Erickson 2015, 27.

⁵⁵ Hawkins 2004. Leaders seek to maximize their personal “esteem.” Finnemore and Sikkink 1998, 898. Research on status from economics, sociology, and psychology shows that people are motivated by status. Harsanyi 1966, Blader and Chen 2012, Ridgeway 2013. For a discussion see Kelley and Simmons 2016.

stigma of backwardness.”⁵⁶ Notwithstanding those who actually desire a negative reputation for unpredictability or toughness, if leaders are sensitive to the international status markers of their state or to domestic legitimacy concerns, these might activate psychological desires for a positive self-image and for social approval.⁵⁷ Enhanced status has psychological benefits and elites may value these as part of their identity and fear being ostracized.⁵⁸ Johnston argues, “The most important microprocess of social influence ... is the desire to maximize status, honor, prestige – diffuse reputation or image – and the desire to avoid a loss of status, shaming, or humiliation and other social sanctions.”⁵⁹

Thus, states and their elites may value their broader reputation for instrumental reasons and “strategically [adopt] popular policies out of social concern for their international reputations rather than out of any existing practice or norm internalization.”⁶⁰ This also means that governments or officials do not need to be persuaded or have internalized the norms and standards; all that matters is that they believe that others will view them as failing with respect to these norms, that *others* ascribe validity to these.⁶¹ Political elites may become “trapped” by the prevailing rhetoric, unable to “craft a socially sustainable rebuttal,” and therefore “compelled to endorse a stance they would otherwise reject.”⁶²

In sum, states and their elites may care about their reputation for both normative and instrumental reasons. Such motivations are often both

⁵⁶ Weyland 2009, 33. For example, sociological intuitionists see policymakers as striving to enhance their international status by borrowing policy solutions from abroad to demonstrate their modernity. Meyer and Rowan 1977, DiMaggio and Powell 1983.

⁵⁷ See discussion by Johnston of the appropriate level of analysis when considering efforts of social influences. Johnston 2008, 95–99. See also Lumsdaine 1993, Shannon 2000, Efrat 2009. Even scholars who question the role of reputation for states stress that individual leaders worry about “their own reputations and status.” Dafeo et al. 2014, 381. Young argues that like private individuals, policymakers “are sensitive to the social opprobrium that accompanies violations of widely accepted behavioral prescriptions.” Young 1992, 176–177. If they believe in the norms but find their own behavior at odds with them, they may also dislike experiencing “cognitive dissonance.” Festinger 1962.

⁵⁸ Franck 1990, 32, 58, Lutz and Sikkink 2000, Johnston 2008, 76, 84.

⁵⁹ Johnston 2001, 500.

⁶⁰ Erickson’s discussion of arms transfer policies provides an example. Erickson 2015, 25.

⁶¹ The sociologist Max Weber acknowledged as much. Jackson 2002, 449. For a discussion of “social reputation,” that makes similar points, see Erickson 2015. Elites need not be persuaded; they may simply strive for public conformity to protect their reputation. Johnston discusses social influence versus mimicry versus persuasion in his book, *Social States*, in which he quotes from Leon Festinger that it is possible to have “[p]ublic conformity without private acceptance.” See Johnston 2008, 24–26. See also Kelman 1958.

⁶² Schimmelfennig 2001, Krebs and Jackson 2007, 36.

compatible and indistinguishable, depending on how one defines costs and benefits.⁶³ They likely operate in concert, either because individual policymakers respond to a combination of these, or because different policymakers in the same country respond differently. In either case, the values states place on their reputations are likely heterogeneous.⁶⁴ Indeed, the sources of concern may not fit as neatly into theoretical boxes as one might suppose. For example, do elites worry about possible consequences like sanctions because they fear losing the funding, or because sanctions are shameful? Some research suggests that sanctions can be used symbolically to influence states, and Arab states have done so at times.⁶⁵ Minor sanctions might be financially bearable; yet the stigma might be unacceptable. Similarly, cognitive dissonance may not just be about emotional stability; individuals may be concerned that others will spot and punish their hypocrisy.

For some reason, however, even scholars who acknowledge that intrinsic and instrumental logics are intertwined default to the instrumental logic and argue “what is at issue is the extent to which these mechanisms lead behavior to deviate from the ideal instrumental policy.”⁶⁶ The burden is always to show that something is not instrumental. From a policy perspective, however, what is at stake is when behavior deviates from norms.

Although these sources of concern can’t all be disentangled, this book will argue and document that reputational concerns are not limited to immediate credibility-based, material repercussions; they are also social and image-based in the sense that states care about being seen as in conformance with the norms in the society of states. This is important for scorecard diplomacy because it means that reputation can be elicited as a tool of influence with or without material leverage, and that it can be elicited across a broad range of issues, not only – as argued by others – “most in security and trade.”⁶⁷

⁶³ Dafoe et al. 2014. For a thoughtful discussion of “social reputation,” that makes similar points, see Erickson 2015. For a similar discussion around the concept of status and war, see Renshon 2016.

⁶⁴ Ragin 2000.

⁶⁵ For a discussion of how “symbolic sanctions” worked in the Arab context during the 1950s, see Barnett 1998. Barnett argues that these worked because states wanted to maintain their dignity, honor, and face.

⁶⁶ Dafoe et al. 2014, 383.

⁶⁷ Downs and Jones 2002, 5112.

WHEN DO REPUTATIONAL CONCERNS OPERATE?
A SIMPLE MODEL

When do broad reputational concerns arise and translate into policy? The answer to this question may vary with the nature of the issue. Conflict-focused research argues that reputational concerns vary with cultural values such as honor and standing and with the number and type of separatist movements or other challengers.⁶⁸ But reputations on war and resolve can often be identified with singular decisions that are likely attached to individual leaders and entail direct interstate interaction, as is also the case in international lending and repayment, for example.⁶⁹ This contrasts with issues such as human rights or environmental performance, which are more diffuse and focus on broader policies, performance, and implementation over time that often involves many actors. Here I present a simple model of three factors that the arguments above suggest will interact to generate reputational concerns on such broader issues and whether these then produce behavioral changes. These factors are *sensitivity* and *exposure*, which together drive the level of concern, and *prioritization*, which drives the translation of concern into policies, as discussed below.

Sensitivity

Reputational concerns depend on how sensitive a government is to any reputational pressures. Sensitivity depends on the salience of the possible practical stakes of low performance, both for states and political elites, as well how much the government sees itself as being in conflict with existing norms and expectations.

One source of sensitivity is what one might call *instrumental salience*: to the extent that reputational concerns are connected to the desire to obtain certain goals, their activation will depend partly on the salience of such goals. These are usually some form of international benefit such as aid, trade, tourism, and office holding in international institutions, or membership in international institutions. For example, when countries want cooperation from the international community, they seek external legitimacy more, as evident when Mexico was negotiating for the North American Free Trade Agreement and President Carlos Salinas was campaigning for the World Trade Organization presidency.⁷⁰ Thus,

⁶⁸ For a discussion of this literature, see Dafoe et al. 2014.

⁶⁹ For a theory on reputation and international debt see Tomz 2007b.

⁷⁰ Kelley 2012, Ch. 2.

when such benefits are more salient, states will be more sensitive to the reputational damage that can block access to these benefits.⁷¹

The extent to which individual elites face consequences may also vary. The more likely officials are to be held directly responsible for poor performance, they more likely they will be to worry about their personal careers. Policymakers may also worry that they or their party might be blamed for underperforming. Bureaucrats may worry about being singled out for inaction or mishandling.⁷² Officials may worry about such practical consequences, even if citizens are not protesting. Rather, they may worry simply if they anticipate criticism.⁷³ If they fear that bad publicity can produce protest or lower their public approval ratings or job evaluation, they will be more sensitive.⁷⁴

Complementing instrumental salience, states will also be more sensitive to reputation if the underlying norms are salient. The *normative salience* depends foremost on how well established the international norms are. Norms encoded in legally binding treaties that most states have signed will be most compelling, whereas the least compelling will be those where considerable contestation remains. Some “norms” are not so much moral in character as ideological. For example, the international community is less likely to have consensus on business regulations than on torture. Even if a norm or expectation is well established, the normative salience of a given state also depends on whether the state and governmental elites identify with these. Certain countries may be comfortable deviating from some norms, even if they are well accepted internationally. For example, some countries may have different views on women’s rights that the state justifies based on religious grounds. Finally, the normative salience will depend on the extent to which actual behavior diverges from expectations. To return to Figure 2.1, the salience of norms depends on the gap between the domestic and international ideal points, and greater discrepancy between these and actual behavior stokes reputational concerns.

Public international commitments can be important in this context. If a state has committed to an international treaty that embodies a set of norms, its declared ideal point is close international norms that outside actors might invoke.⁷⁵ This makes the discrepancy between performance and the standards more damaging. On the other hand, if domestic norms

⁷¹ Keck and Sikkink 1998, 208, Risse et al. 1999, 13–14, Risse and Ropp 2013, 20–21.

⁷² Batson 1987.

⁷³ Cook et al. 1983, Mutz 1998.

⁷⁴ Hendrix and Wong 2013.

⁷⁵ On the importance of public commitment, see Lutz and Sikkink 2000, 657.

differ markedly from international ones, the state will not be as harmed, at least domestically, for failing to live up to standards to which it does not subscribe. This decreases the sensitivity to reputational concerns.

This may lead one to conclude that democratic states are more likely to worry about their reputations because their domestic ideal is likely closer to the international ideal, or because citizens can protest more easily. However, this may vary across issue areas. Domestic ideal points can align well with international ideal points, even in autocracies. This may be especially true if the issue concerns state competencies to deliver services and protect its citizens, to which citizens of autocracies may feel as entitled or with which autocracies may mollify its citizens, a point that aligns well with Huntington's notion of "negative legitimacy." Pollution, for example, has become a huge domestic liability issue for China, around which domestic protests have festered. On the other hand, even democracies may veer from international norms. For example, while Australia has been criticized by the international community for its policy to turn back refugees at sea, the mood at home is one of agreement: many people do not want the large influx of refugees and the government has successfully stoked this sentiment to avoid domestic reputational damage for its non-compliance with international norms.⁷⁶

Normative salience also depends on the relationships between the target and the source of the norms. Because reputation is about legitimacy and identity, criticisms from a desired in-group carry more weight.⁷⁷ If states see themselves as having a certain identity, then they are sensitive to status markers from others who share this identity.⁷⁸ Thus, reputational concerns are likely greater when affinity to the actors eliciting the concerns is greater. Relatedly, because reputational concerns arise when information discredits performance, the authority of the source is important. Those advancing norms must be credible and trustworthy and must exemplify the norms they promote, which imbues them with what some call "normative power" to shape conceptions of norms.⁷⁹ The same might be true if the country has strained political relations with the source, in which case it's easier to dismiss the criticisms as politically motivated. That noted, active efforts to discredit the source signal concern about

⁷⁶ Maguire 2016.

⁷⁷ See Johnston 2001, 2008. Social psychology research finds that people may adopt behaviors if they are associated with the desired in-group. Kelman 1958.

⁷⁸ For a discussion of identity and how it relates to cooperative behavior, see Johnston 2008, 74–75 and Ch. 3 generally.

⁷⁹ On trust as a filter for processing information, see Rieh 2002, Espeland and Sauder 2007. On the value of exemplary behavior, see Sikkink 1993, Monks and Ehrenberg 1999. On normative power, see Manners 2002.

reputational damage, underscoring the presence and weight of such concerns.

Finally, it may matter how good a country's overall reputation is. For example, Norway is able to deviate from the international consensus on whaling without too much damage, because it has a lot of "reputational capital" to draw on. It is well regarded and can afford a light dent in its reputation on this issue. Still, reputations, particularly on issues like debt repayment or peacefulness, can easily be squandered, especially if they are accompanied by leadership changes so that others might infer that a structural change has occurred.⁸⁰

Exposure

Reputational concerns depend not only on how sensitive states are on the issue, but also on the availability and clarity of information and on the exposure that the relevant information brings about. The public element is crucial – reputation depends critically on "the extent to which a behavior is publicly observed."⁸¹ States may be missing the mark but suffer no reputational consequences if nobody knows about it. Clear information is essential for domestic and international actors to evaluate performance and pressure governments.⁸² Exposure is the degree to which a country's behavior, "B," is actively brought into question. Exposure depends on the abundance and clarity of the information as well as on the attention by third parties and the media and focal events.

Active third parties such as NGOs or IGOs can bring attention to the issue in the media and in public discussion.⁸³ A large literature stresses the role of civil society in holding domestic governments accountable. Domestic interest groups may care about the international norms referenced and pressure governments to comply.⁸⁴ Societies with active civil societies may therefore be more susceptible to reputational concerns.⁸⁵

Similarly, large events in a country that focus attention on an issue might also increase the exposure of government officials to the criticisms and heighten the costs of inaction. Media play a particularly critical role.

⁸⁰ For a discussion in the context of debt repayments, see Tomz 2007b.

⁸¹ Dafoe et al. 2014, 376.

⁸² Keohane 1998, Erickson 2015, 30, Guzman 2002.

⁸³ For a discussion of the role of civil society in exposing reputational gaps in state behaviors on arms transfers, see Erickson 2015, Ch. 3.

⁸⁴ Young 1979, Schachter 1991, Sikkink 1993, Jacobson and Weiss 1997.

⁸⁵ Simmons 2009, Erickson 2015, 19.

Even the threat of bad publicity may worry officials.⁸⁶ Underperformance or scandals can attract outsized attention.

Finally, if the source is unreliable or inconsistent, exposure will be diminished, and governments can dismiss criticism more easily and therefore lessen any reputational damage.⁸⁷

Prioritization

When states are sensitive and their performance exposed, they are likely to worry about their reputations. But whether these concerns translate into action depends on their prioritization. Even if a government is concerned about its reputation on a given issue, it may not be able to pay attention or change its behavior. Like others, government officials have limited capacity for attention and action. Agenda-setting theory suggests that when the agenda is crowded with high priority items – for example, all-consuming crime or drug problems as in Honduras – it’s harder to generate attention for other issues.⁸⁸

Furthermore, the individual agency of elites may influence prioritization. Elites may incur personal costs to change. Local issue champions who can act as “sympathetic interlocutors” can help rally domestic reformers and channel the exchange of information and ideas, but if they are absent or irregular, this can stifle progress.⁸⁹ If reforms could interfere with practices from which officials benefit, such as official corruption, opposition will be high, so even if reputational concerns are present, they may be squashed by such corruption. On the other hand, reform-minded elites – be they bureaucrats, government officials, or opposition politicians – can boost attention to the issue and any international criticisms to highlight the reputational concerns within the country. The more such bureaucrats worry and are in a position to make things happen in government, the more their concern matters, as has been pointed out by the literature on the importance of reform-minded teams within countries.⁹⁰

A country’s ability to address an issue will also be affected by its stability: distraction by internal or international wars or uprisings, or political turmoil, can undermine the attention of any government to address

⁸⁶ Mutz 1998.

⁸⁷ Guzman 2002, 1863.

⁸⁸ Kingdon and Thurber 1984.

⁸⁹ Corrales 2006, Chwieroth 2009, Parks 2014. Note that Alter similarly speaks of “compliance partners.” Alter 2014.

⁹⁰ Haas et al. 1993, Jacobson and Weiss 1997, Corrales 2006, Chwieroth 2009, Parks 2014.

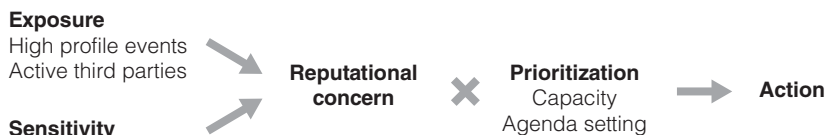


FIGURE 2.2. Factors that influence reputational concern and its translation into action.

less pressing issues. Political instability may also undermine the personal stakes of elites who face uncertain futures and thus may be less motivated to advance their course. When time horizons shrink, investments in reputations are likely to be less valued.

Capacity is also crucial. Even if reputational concerns bring an issue on the agenda, lack of capacity to design or implement change can end good intentions. Such capacity-deficits are known to hamper international reform efforts as well as compliance with international human rights and environmental treaties.⁹¹

Figure 2.2 illustrates how sensitivity and exposure influence reputational concerns, and how prioritization mediates whether these concerns translate into action.

HOW DOES SCORECARD DIPLOMACY ELICIT REPUTATIONAL CONCERNS?

Scorecard diplomacy shares some traits with international institutions, which embody norms and expectations and often provide mechanisms for information sharing.⁹² Through these interactions, institutions can stoke reputational concerns.⁹³ Scorecard diplomacy operates as a type of quasi-institution that promotes information sharing and accountability, and facilitates comparisons of state performance. Like international institutions, scorecard diplomacy can set norms and expectations and provide transparency between citizens and their governments, as well as relative status markers between states. New information, such as indicators or

⁹¹ Haas et al. 1993, Weiss and Jacobson 2000, Kelley 2012, Risse and Sikkink 2013.

⁹² In international relations, information has long been thought to influence the cooperation among states and so international organizations have often gathered and disseminated information about state performance as a tool of influence. Keohane 1984.

⁹³ For the importance of institutions as social environments see Johnston 2001, 2008, Hurd 2008, Erickson 2015.

rating and rankings, alters the base that actors use to assess performance.⁹⁴ Since scorecard diplomacy often evaluates states' performance or criticizes governments, it is well suited to elicit reactions from civil society, media and the international community.

As discussed in the introduction, the scorecard diplomacy cycle incorporates several practical features: monitoring and grading, diplomacy and practical assistance, and third party pressures. The next section discusses how these tap into and elicit concerns about reputation.

The Role of Monitoring and Grading

Scorecard diplomacy engages reputations by using public monitoring and assessment. Importantly, it strategically packages and distributes information in a way that invokes relevant norms, and invites comparisons and judgment. It holds states to some behavior that is simultaneously promoted as ideal.

The grades are crucial. Scorecard diplomacy literally assigns periodic, and highly comparable, performance scores. The scores seek to boil down the distance between the international ideal point, "I," and the state behavior, "B," down to a grade. They thus become a direct representation of the reputational gap. Numbers, as often used for scorecard diplomacy, are frequently treated as authoritative and lodge themselves in people's minds.⁹⁵ Because grades are easily understood representations of government performance and the gap with international standards, politicians may worry that a bad rating can lower their citizens' confidence in their government's legitimacy and thus its ability or right to rule.⁹⁶ Conversely, if a government suffers from poor legitimacy at home, positive international rankings can be a sign of "good housekeeping, which will add luster to the moral and political credentials" of the government's reform efforts, as the chief of staff of the Philippine president said in 2009 about the MCC eligibility criteria.⁹⁷

Grades, ratings, and rankings are potent symbols that shape perceptions about performance. By reducing complexity they designate a preferred interpretation as meaningful. Symbols such as grades are powerful,

⁹⁴ Merry 2011, Davis et al. 2012a, 2012b, 2015.

⁹⁵ Yalch and Elmore-Yalch 1984, Baesler and Burgoon 1994. For more on the power of numbers, see Robson 1992, Andreas and Greenhill 2010, Hansen and Mühlen-Schulte 2012, Kelley and Simmons 2016, 457.

⁹⁶ Finnemore and Sikkink 1998, 903, Hawkins 2002, 32.

⁹⁷ Ermita 2009. Reported in Parks. Parks 2014.

because, as sociologist Beaulieu argued, humans are engaged in “symbolic struggles over the perception of the social world” and they exercise power by trying “to transform categories of perception and appreciation of the social world, the cognitive and evaluative structures through which it is constructed.”⁹⁸ Thus, the use of symbols such as grades to name and label is a political effort to shape perceptions of reality and powerful actors can use such symbols to impose their perspectives.⁹⁹ By categorizing, grades thus socially construct reference points for normative standards of appropriate conduct and encourage others to gauge state behavior against the reference norms. In this way, grades can stimulate the reputational concerns of states and shape perceptions of interest.¹⁰⁰

In addition to their symbolism, their comparative nature makes grades particularly well suited to elicit reputational concerns. Comparisons can shape and reinforce social identities.¹⁰¹ States want to belong to a community, but they also seek to place as high as possible in a social hierarchy, which is why rankings are particularly status-oriented.¹⁰² Thus, states are not only concerned with how they are viewed by other states, but also how they compare to other states. Grades can introduce a competitive element. When states are grouped they can compare themselves with others – a fundamental status exercise. Citizens may not understand the issue, but they understand that their country has been grouped with pariah countries, or has been put on a blacklist of sorts, or is performing worse than similar countries. Thus, by putting the “scores” front and center, scorecard diplomacy engages exactly the kind of out-group feeling that is foundational to theories of shaming.¹⁰³ Grades thus allow governments to reference their state’s standing within a certain category of

⁹⁸ Bourdieu 1989, 20.

⁹⁹ Bourdieu 1989, 22. Eagleston-Pierce draws on Bourdieu’s notion of symbolic power to explain how relatively weaker actors can frame a situation to enhance their position in bargaining with the WTO. Eagleston-Pierce 2013. In *Economic Statecraft*, Baldwin references Harold Lasswell’s work *Politics: Who Gets What When?* which referred to four different types of influence techniques, one of which was “words” or information, or symbolic means, also sometimes called propaganda. Lasswell 1958. Based on this Baldwin defines propaganda as “influence attempts relying primarily on the deliberate manipulation of symbols.” Baldwin 1985, 13. For a discussion of the politics of numbers as symbols, see also Broome and Quirk 2015.

¹⁰⁰ For a discussion of normative approaches to compliance, see Simmons 1998.

¹⁰¹ The sociological concept of “commensuration” captures how actors make sense of the world by grouping and judging objects or entities. For further discussion see Espeland and Stevens 1998, Espeland and Sauder 2007.

¹⁰² For a discussion of this, see Erickson 2015, 27.

¹⁰³ Risse and Sikkink 1999, 15.

states.¹⁰⁴ States might worry about their peer group for various reasons. For example, investors or other actors may use a country's perceived peer groups to make decisions about investment.¹⁰⁵ Grades therefore allow comparisons that stimulate reputational concerns.

Furthermore, the recurrent nature of scorecard diplomacy – assessments are usually issued annually – reveals whether countries are on the right trajectory or whether performance is deteriorating. This iteration engages the country's reputation both relative to others and to itself over time. Crucially, iteration activates anticipation. A country never actually has to be rated poorly to experience concerns about reputation; countries respond to the anticipation of another round of rating or ranking. Thus, similar to a straight “A” student concerned not to get a “B,” scorecard diplomacy can stimulate *status maintenance* behavior: once a country achieves a good rating, it wants to maintain it.¹⁰⁶ For example, when she was told in February 2005 that failure to pass an anti-TIP law could affect Ghana's Tier 1 status, Ghana's Minister of Women and Children Affairs exclaimed, “[W]e must keep Tier 1.”¹⁰⁷ Reiterative and public monitoring and grading is thus a form of systematic “deployment of normative information,” which creates an environment of continuous accountability.¹⁰⁸

Finally, the monitoring aspect is important. Monitoring signals social importance.¹⁰⁹ The philosopher Michel Foucault famously characterized monitoring as a form of control and thus power.¹¹⁰ Research has also identified the famous “Hawthorne effect,” when individuals change their behavior in response to being aware of being observed, akin to what sociologists call reflectivity.¹¹¹ Election monitoring, for example, can improve the conduct of elections without any formal sanctions.¹¹²

In sum, monitoring and grading have unique abilities to elicit reputational concerns.

¹⁰⁴ Checkel 1998, 902, Finnemore and Sikkink 1998.

¹⁰⁵ Gray 2013, Brooks et al. 2015.

¹⁰⁶ Kelley and Simmons 2015, 65.

¹⁰⁷ 05ACCRA364_a.

¹⁰⁸ Kelley and Simmons 2015, 56.

¹⁰⁹ Larson and Callahan 1990.

¹¹⁰ Foucault 1995, 201–202.

¹¹¹ Adair 1984. Kelley and Simmons note, “Sociologists use the concept of *reflectivity* – the tendency for people to change their behavior in response to being evaluated – to explain the effect, for example, of US News and World Report rankings on university priorities.” Kelley and Simmons 2014. See also Espeland and Sauder 2007.

¹¹² Kelley 2012.

The Role of Diplomacy

While ratings and rankings schemes differ in their degree of interaction with the targets, a fundamental feature of scorecard diplomacy is its engagement on the ground. Scorecard diplomacy takes advantage of the fact that the initial grading increases concerns, which makes policymakers more receptive to other efforts. The second step in the scorecard diplomacy cycle is therefore the ongoing diplomacy and practical assistance. Diplomacy can increase reputational concerns, because it gives international actors a chance to share information, to educate local elites, and to underscore the importance of the international norms. The reporting and meetings launch conversations with policymakers about how to define and frame the problem.

Individual level interaction is important. Even Morgenthau recognized the value of interacting with responsible officials who can be identified and held accountable.¹¹³ As discussed earlier, drivers of reputational concerns such as the desire to belong to the in-group or fear of opprobrium operate foremost at the individual level. More recently political scientists have emphasized the interpersonal elements of social influence and socialization.¹¹⁴ Borrowing from the neuropsychology, a study of face-to-face diplomacy suggests it may stimulate unique neurocognitive processes, because “[d]uring face-to-face interaction we move from private to shared experiences.”¹¹⁵ This also aligns with literature in social psychology and sociology, which argues that shame, in particular, is felt in face-to-face encounters, leading people to *avoid* potentially embarrassing encounters.¹¹⁶

New information shared in meetings and reports can influence policymakers’ causal beliefs about the roots of the problem or their understanding of its scope or possible solutions.¹¹⁷ This can lessen the gap between the international and the domestic ideal point and thus increase the normative salience of outside criticisms.¹¹⁸ The provision of new data and

¹¹³ Morgenthau 1950, 189.

¹¹⁴ Johnston 2001, Checkel 2005. Erickson also notes this point. Erickson 2015, 25.

¹¹⁵ Holmes 2013, 839.

¹¹⁶ Scheff and Retzinger 2000, Rutten 2006, 355.

¹¹⁷ This can happen even if policymakers do not change their overall objectives or if they are only self-interested. Goldstein and Keohane 1993, 13. Deutsch 1963, 92 defines simple learning as learning that does not change the goals of actors, but educates them about the state of the world, so that they may choose different ways of accomplishing their goals. This resembles Levy’s notion of causal learning (Levy 1994) and learning in epistemic communities (Haas 1992). Grobe 2010 calls this functional learning.

¹¹⁸ Colonizers, for example, came to see empires as inconsistent with their values of democracy and equality. Jackson 1993.

statistics can motivate states to act when they learn the scope of a problem, something that occurred in connection with the campaign against landmines.¹¹⁹ Scorecard diplomacy may also shape views by requesting information from the domestic government. Research has found that engaging individuals directly in active data collection, processing, and dissemination can shape their cognitive framework,¹²⁰ in effect priming the international norms in local minds and thus increasing attention to performance on those norms. If elites change their beliefs and internalize new ones, that can be considered a form of socialization.¹²¹ That said, beliefs are difficult to observe, and in any case not necessary for the dynamic to operate. Policymakers may simply realize that others view the issue as important and seek to behave more appropriately. This essentially moves the domestic ideal point “D” closer to the international ideal point “I,” which may increase reputational concerns as the gap between “I” and behavior “B” then grows.

Policymakers are particularly likely to seek information about policy alternatives because these often require detailed knowledge.¹²² Policymakers surveyed about their attention to performance assessments mention such information as far more influential than financial incentives.¹²³ The availability heuristic also suggests that policymakers are more likely to adopt solutions that are placed on their radar.¹²⁴ This is perhaps why studies of diffusion have shown that countries often look to similar countries when designing and advocating for new policies at home.¹²⁵ Thus, diplomatic interaction that highlights policy solutions can increase attention to the problem and make policymakers more receptive to proposed solutions.

¹¹⁹ Price 1998, 322. A study on pension reforms in Latin America found that the World Bank influenced reforms more through the provision of information than through conditionality. The Bank provided information that had escaped the attention of policymakers, who, unable to process all relevant information fully, allowed the Bank to influence policy choices, or perhaps was able to use the Bank’s suggestions to push for change. Thus the Bank had more influence through knowledge provision than through conditionality. Weyland 2009, 126–127.

¹²⁰ Von Bogdandy and Goldmann 2008, 242.

¹²¹ Checkel 2001, Johnston 2001, Checkel 2005. This resembles what some call argumentative learning (Risse 2000, Deitelhoff 2009) and Grobe calls sincere learning (2010). This requires “value-based rationale,” more so than scientific-based evidence. Adler 1992, 3.

¹²² Kingdon and Thurber 1984.

¹²³ Parks et al. 2015, 73.

¹²⁴ Weyland 2009.

¹²⁵ Linos 2013.

Diplomatic interaction and repeated requests for information may also stimulate new bureaucratic structures or habits that institutionalize attention to the issue area and facilitate ongoing learning about the problem and possible solutions. New administrative structures, in turn, can transform the capabilities of the states and foster bureaucratic expertise.¹²⁶ This matters, because new institutions can routinize behaviors and change domestic norms. Max Weber argued that habituation was an important mechanism for shaping behavior.¹²⁷ More common behaviors can become routinized because bureaucrats and politicians are “habit-driven actors.”¹²⁸

In sum, the individual diplomacy calls attention to the issue and shapes local views and ideal points. The more this occurs, the more concerned elites become with the issue and their related reputation.

The Role of Practical Assistance

Not all ratings and rankings are connected to extensive assistance, but those that are will be more likely to work. By funding the government or NGOs or IGOs, scorecard diplomacy can contribute to structures that might shape local norms and practices and thereby increase reputational pressures to live up to those standards. Funding for capacity building or training may introduce new solutions to problems and normalize certain behaviors. Even if the assistance itself may not be effective, it can empower recipients and broaden the set of voices that participate in the national discussion.¹²⁹ Diplomacy and assistance extend attention to the issue far beyond the peak attention it might receive surrounding the release of the report itself and provide ongoing opportunities for interaction, institution building and information transfers.¹³⁰ Thus, while structures are often thought to prevent change,¹³¹ new exogenously introduced structures can explain change in local norms and practices.¹³²

Practical assistance or aid can also increase the *instrumental salience* of the grades. If ratings or rankings are linked to aid or other economic

¹²⁶ Pierson 1993, 58, Skocpol 1995.

¹²⁷ Weber 1968 [1925], Hopf 2010.

¹²⁸ Rosenau 1986, 861–870.

¹²⁹ While the extensive literature on building and technical assistance is mixed with respect to results, research suggests that targeted assistance like this is more likely work than general foreign assistance. Collins 2009, 371–381, Scott and Steele 2011, 53.

¹³⁰ Risse and Sikkink 1999, Johnston 2001, Checkel 2005.

¹³¹ See discussion in Hopf 2010 and references to Bourdieu 1989.

¹³² Neumann 2007, Hopf 2010.

advantages, this resembles standard uses of conditionality or sanctions. Because a strong country like the US can obstruct the flow of economic benefits to other countries, elites may fear economic fallout for their country. The linkage of the Special 301 Report and US trade, for example, generated concerns about costs to trade with the US.¹³³ That said, scholars have long questioned the efficacy of direct economic leverage in achieving policy changes¹³⁴ and criticized them as impracticable and harmful.¹³⁵ Some successes nonetheless exist: Sanctions were widely credited with reforms in South Africa, some research has found that economic pressure can destabilize leaders, the European Union (EU) was able to incentivize reforms in candidate states by linking them to membership, and the US inclusion of human rights clauses in preferential trade agreements has been effective.¹³⁶ More recent work finds that *ex post* performance-based incentives such as the Millennium Challenge Corporation's (MCC) use of performance indicators to award aid can induce reforms, although this so-called MCC effect may wear off over time if the conditionality is implemented inconsistently – a lesson that might well also apply to scorecard diplomacy more generally.¹³⁷ Thus, targeted economic pressure that credibly links policies with economic consequences should matter.¹³⁸

Because the assessments are repeated they may also produce anticipatory effects similar to the “hidden hand” of sanctions: the very threat of punishment may lead to efforts to avoid them.¹³⁹ If the rating is linked to punishments or rewards then the recurrent assessments represent an ever-present possibility of economic consequences. Even well performing countries fear the potential consequences of downgrades, a dynamic that has been played out in the Special 301 reports with their more concrete trade implication.¹⁴⁰

¹³³ Newby 1995.

¹³⁴ Collier 1997, Killick 1997.

¹³⁵ Hufbauer et al. 1990.

¹³⁶ On South Africa, see Klotz 1995. On destabilizing leaders, Marinov 2005. For discussion of EU incentives for reform, see Schimmelfennig et al. 2003, Jacoby 2004, Kelley 2004b, Vachudova 2005. Finally, for the use of human rights conditions in trade, see Hafner-Burton 2005.

¹³⁷ On the inducement of reforms, see Parks and Rice 2013. On the fading of the MCC effect see Öhler et al. 2012.

¹³⁸ On the importance of credibility see Kelley 2004a, 51, 190–191.

¹³⁹ Drezner 2003. This also relates to anticipatory compliance effects of trade negotiations, which suggest that states change their behaviors before agreements enter into effect. Kim 2012, Dür et al. 2014.

¹⁴⁰ Newby 1995, 36.

Even if the scorecard linkages to aid or trade are not explicit, elites may still worry about economic consequences of a poor reputation in a given area.¹⁴¹ This may be true even in the human rights area, where shaming by international NGOs has been linked to lost investment for targeted states.¹⁴² In Thailand, for example, the business community was concerned that the 2014 Tier 3 rating would harm trade, especially the fishing industry. This continued even after President Obama waived the sanctions threat.¹⁴³

The Role of Third Parties

Scorecards are designed to engage third parties and increase *exposure*. Because reputation is relative to existing norms and multiple audiences, the attitudes of other parties are important. Third parties can amplify reputational concerns because they reinforce the larger global set of norms and standards. Once scores are public, other actors such as NGOs, IGOs, and the media augment the visibility of the scores and thus increase exposure and subsequent reputational pressure. This can also trigger market pressures of various kinds. Third parties are not necessary to the success of scorecard diplomacy, but they boost the odds.

The media disseminates ratings or tiers easily, which increases the exposure that activates reputational concerns. For example, the media coverage following the US Special 301 reports about intellectual property rights has been described as media hysteria. Ratings are much simpler than narratives,¹⁴⁴ and thus easier to use as “psychological rules of thumb”¹⁴⁵ and to present in media stories. Importantly, media coverage can worry policymakers even if citizens don’t take to the streets. Even the anticipation of publicity and negative domestic reactions can pressure government officials.¹⁴⁶ Research has shown that elites in general pay much more attention to media than the general public, because they believe that the media does, in fact, affect public opinion.¹⁴⁷ Thus, negative media coverage can create pressure if it creates the *expectation* of

¹⁴¹ Lebovic and Voeten 2009.

¹⁴² Barry et al. 2012. Research also suggests that investors care more about appearances (as in having ratified treaties) than actual human rights performance, which suggests that global ratings and rankings should matter. Garriga 2016.

¹⁴³ Charoensuthipan 2014a, 2014b.

¹⁴⁴ Espeland and Stevens 1998, 316, Löwenheim 2008, 257–258.

¹⁴⁵ Sinclair 2005, 52.

¹⁴⁶ Cook et al. 1983.

¹⁴⁷ Mutz 1998.

protest or of lowering public approval.¹⁴⁸ Of course, domestic officials can also use the media to try to save face by rebutting or rejecting the content of the public scorecard. Indeed, it's noteworthy if domestic elites feel pressured to do so in the first place. Governments are also not always successful at controlling the flow of information, and cannot control the international press, much of which now reaches citizens through the Internet. Thus, media, both international and domestic, can increase the reputational concerns of domestic elites.

IGOs and NGOs can also increase the exposure to scorecard diplomacy. Funding these actors allows them to carry out programs, and boosts their standing within society.¹⁴⁹ Their programs, in turn, increase attention to the issue and therefore the exposure of the state to any criticisms of its performance. This is the *inverse* of a tactic scholars have called the “boomerang” effect, in which domestic NGOs unable to pressure their own governments directly can harness the power of outsiders who then pressure the government of those NGOs.¹⁵⁰ In the case of scorecard diplomacy, NGOs are not necessarily unable to pressure their own governments, that is, it is not limited to unreceptive authoritarian settings. Furthermore, the pressure is initiated internationally in an ongoing – rather than ad hoc – fashion. Still, a similar dynamic can ensue: NGOs can use poor ratings to augment their pressure on their governments.

The reports may also provide NGOs with more specific substance with which to pressure the government. One global survey found that over half of civil society actors agreed they were empowered to advocate for reforms because the Millennium Challenge Corporation tied aid eligibility to measures of policy performance.¹⁵¹ NGOs can also use scorecard diplomacy the way they mobilize around international legal commitments by using the commitments to hold officials accountable.¹⁵² A study of the Extractive Industries Transparency Initiative (EITI), which declares countries to be in or out of compliance, has found that the EITI “improves the capacity of civil society to hold governments to account, whether governments like it or not.”¹⁵³ Thus NGOs can use scorecard

¹⁴⁸ Hendrix and Wong 2013.

¹⁴⁹ Collins 2009, 371–381, Scott and Steele 2011, 53.

¹⁵⁰ Keck and Sikkink 1998.

¹⁵¹ Parks 2014, 229.

¹⁵² Simmons 2009.

¹⁵³ David-Barrett and Okamura 2013, 3.

diplomacy reports for information, but also directly for advocacy and to engage in conversations about problems and needs, all of which increases the governments' exposure to the reputational concerns.

IGOs can also boost scorecard diplomacy. They have long been argued to play important roles in transnational advocacy.¹⁵⁴ Their work on the ground encourages information exchanges with the creators of scorecard diplomacy: IGOs may become sources for information into the regularized reporting, but they may also become primary consumers – and thus legitimizers – of that information, which increases the weight of the grades in the minds of policymakers. For example, organizations such as the United Nations (UN) and the IOM use the TIP Report as a source of information.¹⁵⁵

Other actors can also increase the consequences of scorecard diplomacy by incorporating scorecards or performance assessments into their own criteria for actions, as the Millennium Challenge Corporation has done for awarding aid. At other times, the link is more direct, for example, if the creators of scorecards fund other actors to engage in efforts that align with the basic goals. In this way, concerns about funding may increase, thus increasing the instrumental salience and, therefore, concerns about reputation.

EMPIRICAL EXPECTATIONS

About the Production of Scorecard Diplomacy

The arguments in this chapter suggest a multitude of hypotheses and observable implications. On any given issue the actual set of these will vary with the available evidence and the context. In general, however, if scorecard diplomacy operates as proposed, then, in an information-rich environment, we should observe evidence of the conduct of diplomacy and the engagement of third parties. We should see active reporting and monitoring connected with diplomacy at meaningful administrative levels, and we should see media coverage of the reports, and involvement of NGOs and IGOs. NGOs, in particular, should use the report to pressure their governments.

Chapters 3 and 4 examine these behaviors in the context of human trafficking.

¹⁵⁴ Keck and Sikkink 1998.

¹⁵⁵ This is evident in their reports. See for example UNODC 2009.

About the Verbal Reactions to Scorecard Diplomacy

Furthermore, if scorecard diplomacy generates concern about ratings as argued, then to the extent that it is possible to observe how countries react, we ought to see that those who fare the worst react more often, and in many instances react negatively. In an effort to minimize the normative salience of the criticism, attempts to discredit the creators of the scores might occur. There may also be contestation over the norms if international and domestic ideal points differ. Countries that get more exposure through media coverage, for example, might also react more.

This chapter has also stressed that countries' reputational concerns are broad, that is, they are both instrumental and normative, and they include concerns about image and standing in the international community. While it may not always be possible to observe, countries might explicitly express such concern about their image or standing, even if they don't face any aid or trade threats as a result of the score, or they may underscore such concerns by explicitly comparing their scores with those of other countries, especially those they consider peers. If the reference to international norms is important for states, we might also see that those that have publicly committed to the international norms underlying the scores might react more often. On the other hand, if countries worry about ratings because of aid or trade, we might see discussions in reaction to poor grades in particular, and countries whose aid is threatened might react more.

Chapter 5 examines these behaviors in the context of human trafficking.

About the Policy Responses to Scorecard Diplomacy

The goal of scorecard diplomacy, of course, is not just to get states riled up, but to get them to change their behavior, whether that be policies, regulations, programs, or whatever the relevant outcomes are. If states care about monitoring and grading, then we should observe that they respond when they are included in the monitoring scheme, when they get harsher grades, and when they are downgraded in their rating or ranking.

How would one know that these patterns are really due to the scorecard diplomacy? Here one can derive further observable implications. For example, since the ratings are supposed to affect change through concerns about reputation, then it might be possible to observe not only that scores relate to the level and nature of verbal reactions, but also that countries that react are more likely to change their behavior. Furthermore,

if states are motivated by the prospect of a rise in their rating, if one has good information, then one might also be able to observe extra efforts by states to accomplish certain goals as the deadline for the next scorecard cycle approaches. Indeed, with very good access to information, one might even be able to observe that officials explicitly state that they are undertaking reforms with the hope of improving their scores.

Chapter 6 explores these behaviors in the context of human trafficking.

About the Factors Modifying the Policy Reactions to Scorecard Diplomacy

Finally, this chapter has presented a simple model for understanding when scorecard diplomacy is more likely to matter. Specifically, policy responses should depend on states' sensitivity, exposure, and prioritization. Therefore, policy responses should be more likely when states are more sensitive, that is, if the country wants to boost its international image, if its norms align with the international norms, or if it faces economic repercussions of a low score. This sensitivity will be heightened the more credible the scorecard creator is. Policy responses should also be more likely when states are more exposed. That is, when third parties are actively attracting attention to the issue, or when local events bring attention to issues, and when scorecard diplomacy is credibly implemented in a country, giving it greater weight. Finally, states are more likely to change their behavior in response to scorecard diplomacy, the more they are able to prioritize the issue, such as when local officials take the issue on, when the government is stable and crisis free, or when domestic actors do not have large stakes in preventing change.

Chapter 7 examines these behaviors in the context of human trafficking.

SUMMARY

This chapter has defined reputation as used in this book, explained its nature, and discussed why states and their governments value a good reputation. It has offered a simple model to explain the factors that modify whether states become concerned about their reputation on a given issue and whether they actually do something about it. Finally, it has considered how the features of scorecard diplomacy stimulate states' concern for their reputation.

The argument takes its starting point in the claim that states are not merely concerned about their reputation in terms of the credibility of

their threats and promises in keeping commitments or their resolve in war; rather, states care about their reputation about their broader performance across many issues as assessed by multiple audiences. A good reputation is important for states' image and identity. Governments know that others – citizens, civil society, and other international and domestic actors – assess their performance relative to a set of standards and norms across a number of issues and that this assessment undergirds their legitimacy and status, both of which bolster their ability to govern and interact with other states.

States sustain a good overall reputation when they adhere to norms more broadly and adopt productive policies, yet, in a world of increasing publicity and accountability, their reputation can suffer even when they falter in a narrow area that attracts attention. Governments therefore worry about their reputations across multiple areas and generally prefer to satisfy the norms and ideals promulgated domestically and internationally. A few exceptions notwithstanding, even authoritarian governments have reasons to worry about their performance. Since the procedural foundation for their authority is weak, to suppress revolt they depend even more on successful performance in other areas.

Performance gaps arise when a state veers from established norms and expectations. How concerned a state becomes about this gap depends on several factors that influence its *sensitivity* and *exposure* to criticism. Sensitivity describes the weight that governments assign to the performance gap. Such gaps are more salient to a state if those promulgating the standards belong to a group with which it identifies and if the norms or ideals are well established. Nonetheless, if the domestic standards differ markedly, the international gap will be less salient and the state will be less concerned. Likewise, a state is more sensitive to performance gaps that might affect other objectives, for example, if the state is seeking some benefit that's tied to the specific issue area. Exposure describes how much credible attention the performance gap gets. This depends both on the authority of the sources of criticisms, as well as the volume and quality of information about the government's performance. Third parties can magnify the exposure.

States' concern about their reputations gives the international community an opportunity to influence them. The use of scorecard diplomacy and similar efforts can garner attention and highlight the performance of a state over time and relative to peers. The repeated nature of this exercise makes countries concerned about current and future ratings. Its public nature lets other actors reinforce the norms, which increases the

government's reputational concerns. The use of monitoring and grading therefore can raise governments' concern about their reputations, especially if the efforts are further supplemented by diplomatic engagement.

Whether concern translates into action depends on the priority the government can give it. States have crowded agendas, experience instability, and may lack capacity. Some states may assign a lower relative value to its reputation on a given issue when the domestic agenda is very crowded. Surely, when leaders gas their people, as has Syria's president Bashar al-Assad, the reputational cost to them is considered tolerable in light of their other objectives. Getting from concern to action thus remains a context-dependent challenge. Nonetheless, without other sure-fire tactics to influence recalcitrant governments, it makes sense to attend to the question of reputation and how to stimulate concerns about performance. In the absence of military or strong coercion, reputational concern is a prerequisite for action, so it's important to understand how it arises and can be stimulated. This book therefore moves on to examine what we can learn about reputation and scorecard diplomacy from the case of human trafficking.