Human Rights Abused? Terrorist Labeling and Individual Reactions to Call to Action

Ana Bracic[1] and Amanda Murdie[2] (iD)

Abstract

What leads individuals to be motivated to act for human rights causes? Human rights organizations (HROs) often use personal and emotional stories of the abused in order to gain the attention of individuals reading newspapers or emails directly from the organization. McEntire, Leiby, and Krain show that personal frames are most successful at increasing knowledge about a specific human rights situation and motivating individuals to act. However, HROs are not operating in a political vacuum; repressive governments often try to spin information about abuses and the abused to their advantage. This study uses an experimental approach to address how the discursive interactions between states and HROs influence individual-level support of HRO efforts. When governments respond to HRO claims by labeling the abused as a terrorist, individuals are much less likely to be spurred into certain types of action, like signing a petition, and feel differently about the specific case. Interestingly, we do not find that trust in the information provided by an HRO is harmed by being associated with someone labeled a terrorist.

Keywords

human rights, human rights organizations, NGOs, experiments

There can be a fine line between defending the rights of terror suspects and defending their views. Amnesty International has been thrust into a very public struggle about where exactly that line is.

—Guz Raz, NPR Host of "All Things Considered" February 27, 2010.

For decades, Amnesty International and other human rights organizations (HROs) have been working for the release of "prisoners of conscience" throughout the world. These individuals are often serving long prison sentences in excruciating conditions; they have been jailed for nothing more than their beliefs and nonviolent advocacy efforts. Once a prisoner of conscience is identified, Amnesty International and other HROs begin a media campaign specifically designed to increase pressure on a regime to release the prisoner. Both directly, through emails and social media, and indirectly, through media coverage of their press releases and reports, HROs try to get individuals around the world to understand the plight of the prisoner of conscience and join efforts to get the prisoner released. For those in the same country as the prisoner, HROs hope that the poignant information they provide will lead citizens to pressure their government for the prisoner's release. HROs also hope that individuals in third-party countries will pressure their own political representatives to increase diplomatic pressure on the repressive regime. Because of this increased attention and diplomacy from a now-empowered transnational advocacy network, regime leaders are often forced to rethink the imprisonment of the identified prisoner (Keck and Sikkink 1998; Risse, Ropp, and Sikkink 1999, 2013).

Although advocating on the behalf of prisoners of conscience is just one small portion of what Amnesty International does, collecting information and sharing stories of abused individuals has been a crucial tactic in the repertoire of many HROs. This tactic is often called "naming and shaming" or "shaming and blaming." Recent cross-national research has found that naming and shaming tactics can influence foreign policy decisions (Murdie and Peksen 2013, 2014), change investment patterns (Barry, Clay, and Flynn 2013), end genocides and

[1]Michigan State University, East Lansing, USA
[2]University of Georgia, Athens, USA

Corresponding Author:
Amanda Murdie, University of Georgia, 327 Candler Hall, Athens, GA 30602, USA.
Email: murdie@uga.edu

politicides (DeMeritt 2012; Krain 2012), and, in certain situations, improve human rights practices (Hendrix and Wong 2013; Murdie and Davis 2012).

The success of HROs and of the overall transnational advocacy network, however, rests on the ability of HROs to influence *individuals*—both elites and non-elites—to take actions to stop human rights abuses. If the information HROs produce fails to motivate individuals into action, the transnational advocacy network breaks down before it can be argued to have any potential impact on repressive regimes.

HROs often use personal and emotional stories of the abused in order to gain the attention of individuals reading newspapers or emails directly from the organization. When compared to informational or motivational frames of abuse, McEntire, Leiby, and Krain (2015b) show that personal frames are most successful at increasing knowledge about a specific human rights situation and motivating individuals to act.

However, HROs are not operating in a political vacuum; repressive governments often try to spin information about abuses and the abused to their advantage (Risse, Ropp, and Sikkink 1999, 2013). Many prisoners of conscience have been accused of violence and labeled terrorists by the governments that imprison them, often without clear evidence that this is actually the case. This information makes it to the same news outlets that are covering the HRO's press releases and reports. News stories on the abuses HROs uncover regularly include counterclaims made by repressive governments. The media may include these counterclaims in an effort to appear unbiased.

In this project, we conduct a survey experiment to examine whether information—coming directly from the government accused of the abuse—claiming that the abused person is a terrorist is enough to stop individuals from supporting a personalist HRO campaign. Does terrorist labeling affect a person's willingness to trust the HRO and their feelings on the case? Does it limit an individual's willingness to support the campaign?

In extant studies of HROs, advocacy networks, and shaming, the counteractions of governments in response to HRO shaming are rarely addressed.[1] This study uses an experimental approach to address how the discursive interactions between states and HROs influence individual-level support of HRO efforts. Individuals often have a front-row seat to this discursive interaction in the press and social media. When a repressive regime labels a prisoner of conscience a terrorist, does this limit the success of an HRO's personalist frame? If so, understanding and anticipating government counteractions to HRO attention is necessary to improve human rights. Moreover, any effects that government counterclaims have on an individual's perception of the trustworthiness

of the HRO itself could have lasting consequences for the success of the organization. Many repressive governments have been increasing their targeting of civil society groups, a phenomenon known as "closing civil society space" (Wilson 2016, 331). If individuals also lose faith in HROs when repressive governments issue counterclaims, this could doubly restrict the ability of HROs to improve human rights.

This paper proceeds as follows. We first provide some background on the work of HROs and the strategic use of information by HROs and governments. We then present our theoretical argument and hypotheses. Our experimental design is then presented and our results are discussed. We conclude by situating our findings into the larger practitioner and academic literature and provide some calls for future research.

HRO Accounts of Government Abuse

HROs are crucial actors in overall transnational advocacy efforts; their attention is commonly seen as the critical factor in both mobilizing international condemnation and strengthening local efforts for human rights improvement (Keck and Sikkink 1998; Risse, Ropp, and Sikkink 1999). This process of human rights improvement, however, is not strictly linear and often involves a very serious back and forth with repressive regimes. Risse, Ropp, and Sikkink (1999) lay out this logic in their discussion of the five phases of potential human rights change, which they term the "spiral model" (Risse and Sikkink 1999, 22). In the first phase, HROs and other transnational advocacy actors become aware of abuse and begin their attempts for more international attention and condemnation, starting with just gathering information about the abuses. In the second phase, as HRO shaming against the targeted state ratchets up, the repressive government's "initial reaction . . . is almost always one of denial" (Risse and Sikkink 1999, 7).

At this second phase of the spiral model, targeted governments often try to frame the HRO and accounts of the abuses as ignoring the security needs of the state; in other words, the repressive state tries to invalidate the claims of the HRO, often with references to the need for order and national security. It is at this stage that claims of violence and terrorism frequently appear, as the repressive state tries to isolate "domestic human rights organization and international pressures by identifying these groups as conscious or unconscious accomplices of terrorism" (Risse and Sikkink 1999, 23). As Risse and Sikkink (1999) point out, these critiques can be damning to the transnational advocacy network and the "transition to the [next] phase [of human rights improvement] constitutes

the biggest challenge for the transnational human rights network" (Risse and Sikkink 1999, 24).

Only with continued pressure and a somewhat vulnerable repressive state does the spiral model transition into the third phase, that of the state deciding to make "tactical concessions" to the movement in order to avoid further attention. At this stage, the HRO can no longer be ignored or dismissed. Governments and HROs begin to establish more of a back and forth dialogue related to the abuses in question; this dialogue is referred to as a "discursive process" (Risse and Sikkink 1999, 14). If domestic pressure continues to build, the final phases of human rights improvement begin, moving the state into first a phase of "prescriptive status," where treaties are signed and norms are strengthened, and then later to the phase of "rule-consistent behavior," where human rights norms are internalized (Risse and Sikkink 1999, 32).

As mentioned, at the national or cross-national level of analysis, there exists much evidence consistent with this theoretical framework. HRO shaming efforts bring international attention and change foreign policy (Barry, Clay, and Flynn 2013; Murdie and Peksen 2013, 2014). They also, at least for certain human rights abuses in certain situations, lead to human rights improvements (DeMeritt 2012; Hendrix and Wong 2013; Krain 2012; Murdie and Davis 2012).

All of this evidence and the whole process of human rights change, however, rest on two critical, and yet often simply assumed, processes: (1) HROs first have to find out about abuses and begin a campaign around the issue and (2) individuals have to be swayed into action by the information provided by HROs. An explosion of recent work addresses the first process, the actual choice by HROs of cases on which to focus (Bob 2005; Carpenter 2007; Hendrix and Wong 2013; Hill, Moore, and Mukherjee 2013; Ron, Ramos, and Rodgers 2005). Much of this work is at the cross-national level, concluding that geo-political concerns influence which countries receive attention (Hafner-Burton and Ron 2013; Ron, Ramos, and Rodgers 2005), as do need and poor human rights conditions (Hill, Moore, and Mukherjee 2013; Ron, Ramos, and Rodgers 2005). Other work examines how issue salience and fit matter for organizations, concluding that particularly poignant cases of clear issues of bodily harm often receive the most attention (Bob 2005; Brysk 2014; Carpenter 2007; Keck and Sikkink 1998). Nonetheless, as Hill, Moore, and Mukherjee (2013) found when focusing on Amnesty International, concerns about credibility keep HROs from exaggerating their claims.

The second oft-assumed process through which HROs work to improve human rights has received far less attention. Once the HRO has decided on a particular campaign issue, HROs need to motivate elites and non-elites, both in the repressive country and internationally.

For elites, HROs want to mobilize international bureaucrats and policymakers to take actions to stop the abuse in the repressive state. Although these actions could be promises of increased foreign aid or economic opportunities if the regime stops the practice, many times these actions are threats of foreign policy actions that will happen if the regime continues the repression. Threats of reduced foreign aid, sanctions, or interventions are intended to make the regime leadership reconsider the utility of the abuse. Motivated international elites could also call for United Nations' attention to the abusive regime; they could demand transitional justice. Domestic elites may put pressure on the regime directly, perhaps wanting to avoid the economic consequences of a boycott or imposed sanctions.

For non-elites outside of the repressive state, HROs want to mobilize individuals to take actions to increase pressure on the regime. These actions include donating to the cause or organization, volunteering their time, calling their representatives, and boycotting products from the abusive regime. For non-elites inside the repressive state, HROs hope that information on the abuses will help in building domestic coalitions against the abusive practices, perhaps leading people to the street to protest against the regime.

How are individuals mobilized to act in ways that pressure an abusive state? Human rights scholars have spent much time empirically examining the foreign policy tools by which international governments and organizations try to sway repressive regimes (Peksen 2009, 2012; Poe 1990; Richards, Gelleny, and Sacko 2001). Far less attention has been devoted to the process by which individuals are motivated to demand this foreign policy action in the first place. At the cross-national level, we do have evidence that efforts by HROs to name and shame specific states change aggregate individual opinions regarding human rights conditions in those states (Ausderan 2014; Davis, Murdie, and Steinmetz 2012). However, there are still many questions about the type of individuals and the type of messages that resonate and mobilize. Do all actions by HROs mobilize in the same way? What would cause an individual to stay unmoved or unaffected when confronted with evidence of abuse? Answering these questions is critical to fully understand the transnational nature of human rights advocacy. Without answering these questions, we are missing a crucial source of variation in the success of the boomerang and spiral models; without swaying individuals to act, the whole transnational advocacy network shuts down.

Two recent studies provide some insights into these questions. Both of these studies rely on experimental research designs. First, Ausderan (2014) provides both an observational and an experimental look into whether HRO shaming influences public opinion on human rights

conditions. Although Ausderan (2014) finds observational evidence that HRO shaming reports influence public opinion on human rights conditions, he finds no statistically significant evidence in the experimental study, which focused on whether information about Amnesty International accusing the United States or India of human rights abuses influenced human rights opinions in those states. Ausderan's (2014, 91) experimental study was conducted on Amazon's Mechanical Turk (MTurk) and had a dependent variable that was a seven-point scale of human rights conditions in the respondent's home country. The lack of findings in this experiment could be associated with the small number of respondents (291 individuals) or could indicate that this particular HRO message was ineffective in the general population.

The second recent study on this issue is McEntire, Leiby, and Krain (2015). As noted, this study focuses not only on the treatment of whether an HRO report is available but also on the type of framing used by HROs. McEntire, Leiby, and Krain (2015) identify three common frames used by HROs: (1) informational frames, where the focus is on "facts and statistics"; (2) personal frames, where emotional stories are provided; and (3) motivational frames, where the role of the potential advocacy success is highlighted (409). McEntire, Leiby, and Krain (2015) offer very nuanced hypotheses for how these frames will influence opinions, mobilization, emotional reactions, and knowledge. Overall, it is personal frames that McEntire, Leiby, and Krain (2015) argue will be most effective: "[f]acts alone should do less to make others focus on the problem as central and in need of their action than efforts that help people identify with other humans in distress" (412). Like, Ausderan (2014), McEntire, Leiby, and Krain (2015) also use an experiment to test their hypotheses over MTurk, this time with 826 respondents. They find much support for their hypotheses, especially concerning the importance of personal frames. Later work by these authors reinforces the importance of personal frames (McEntire, Leiby, and Krain 2017).

This work is essential to our understanding of the process of human rights improvement; in particular, these experimental studies represent critical advances in the field. Although it has long been asserted that evocative frames are necessary for human rights actions, McEntire, Leiby, and Krain (2015) empirically show that personal frames are key to human rights mobilization. Taken together with Ausderan (2014), we now know that information alone may not influence opinions or mobilize actions; it is the personal stories of those harmed by abuses that influence opinions and affect mobilization.

When we connect these important results back to the boomerang and spiral models, however, we see another critical lacuna. HROs are not releasing personal stories in a vacuum. As Risse and Sikkink (1999) make clear, there is a discursive process occurring by which the repressive state tries to discredit the organization and dismiss any abuses as necessary for security. However, no existing work on individual responses to HRO shaming takes the discursive process of HRO–government interactions into account. In other words, none of these important experimental pieces address the fact that individuals are making decisions about whether to support an HRO's cause at the same time that the government can, and often does, issue statements condemning the HRO and the abused. In the news coverage of HRO campaigns, the responses of the government are often reported alongside the statements of HROs. For those getting the HRO's personal frame over email or social media, government responses are readily available and often issued in direct response to an HRO press release or social media statement.

Many accusations against HROs and the individuals they represent can be thought of as part of the normal discursive process of the second phase of the spiral model of human rights improvement. In labeling an individual a terrorist or bringing up questions about the organization, the regime, which has been called out for its use of repression, is trying to both deflect and deny the charges against it. According to the spiral model, through this deflection and denial, the regime hopes that it will avoid further advocacy and attention.

What role, if any, does the government's response have on an individual's likelihood to believe the HRO and support the person identified by an HRO's report? To our knowledge, this question has never been examined with respect to HROs and the information they produce. Addressing this question is critical for connecting the recent advances from the experimental literature to our theoretical understanding of transnational advocacy and human rights improvement.

Government Responses to HRO Accounts of Abuse

As part of the discursive process between the government and advocates, repressive regimes try to both vilify and discredit the HRO bringing attention to an abuse and the abused themselves. Governments often stress how HRO reports are missing the violent actions of insurgents, implying that the HRO's report should not be trusted. For example, in 1982, as Amnesty International and other organizations were condemning mass killings in Guatemala, United States Assistant Secretary of State Thomas O. Enders sent an open letter to Amnesty International, drawing attention to abuses committed by insurgents (Weinraub 1982). A similar situation happened when HROs began reporting on political disappearances

in Argentina; HROs were called "Marxists who did not understand the threat of terrorism" (Bouvard 1994, 84). And, more recently, both U.S. President George W. Bush and Vice President Dick Cheney lobbed criticism at Amnesty International's 2005 report on human rights abuses at Guantanamo Bay (BBC 2005).

The merit of these responses from governments is often questioned. However, one could expect that many potential sympathizers of human rights causes are aware of the common vilification of victims by repressive governments. To these supporters, the charges that governments raise about prisoners of conscience or even abused individuals where there is good evidence of actual participation in terrorism and violence are of little consequence. They would remain resolute in their focus on the government's abuse and the HRO efforts to end it.

However, it seems likely that governments use denial and deflection because, at least in part, this strategy is effective for them. Upon hearing these responses, individuals could question for legitimacy of the organization and change their opinion of the abused. Hortsch (2010) discusses how HROs and other international nongovernmental organization (INGOs) are concerned with their legitimacy and reputation. An organization's authority rests on the veracity of the claims they are making; if the organization is seen as distorting or hiding facts, this can harm their standing (Hortsch 2010). For example, after September 11, HROs have had their authority increasingly questioned for speaking out on behalf of torture victims who have been charged or suspected of terrorism. Amnesty International was thrust into the spotlight in early 2010 when one of its senior officials stepped down after concerns were raised about its partnership with a suspected terrorist-supporter, who wrote about his time at Guantanamo Bay. The issue called into question the difficult decision HROs face when continuing to support those accused of terrorism (NPR 2010).

Government vilification of HROs is part of a growing trend of repression of civil society. Starting in the 1990s and increasing precipitously in the 2000s, repressive regimes have increased restrictive laws and negative rhetoric about civil society organizations, especially international groups involved in human rights and democracy promotion. Anecdotal evidence suggests that public opinion of civil society organizations has soured during the time since the pushback phenomenon began (Carothers and Brechenmacher 2014, Carothers 2016).

Of course, statements by governments in response to HRO shaming are not just directed at HROs, they are also directed at the individuals on behalf of whom the HROs are advocating. Governmental rhetoric often labels even prisoners of conscience as terrorists and violent offenders. Many times, governmental responses go beyond rhetoric, justifying imprisonments with often unsubstantiated charges of violence and/or terrorism. Take, for example, the cases of three Al Jazeera English reporters in Egypt, Peter Greste, Mohamed Fahmy, and Baher Mohamed. All three individuals were detained in Egypt in December 2013 for, as Amnesty International reported, "the peaceful exercise of their right to free expression" (January 29, 2014). The Egyptian government, however, charged them with collaboration with terrorists. Greste, an Australian, was released and left Egypt, but even though the charges were widely regarded as a sham resulting only from reporters meeting with regime's political opponents, Fahmy and Mohamed remained in prison until September 2015. Irish national Ibrahim Halawa, a participant in the Arab Spring protest, was also held in Egypt for over four years, despite human rights advocates labeling him a prisoner of conscience and calling for his release.[2] Similarly, reporter Reeyot Alumu served a five-year sentence in Ethiopia and was "branded a terrorist" for reporting on social issues while not conforming to the regime rhetoric.[3] By labeling these individuals as terrorists, the Egyptian government could have influenced individual opinions of the cases and the likelihood that people would act to help the imprisoned.

Although there are many reasons to expect that discursive responses by governments discredit HROs and limit mobilization on behalf of the abused, this critical step in the spiral model has not been examined. We expect that individuals exposed to the discursive responses of governments will be less likely to support the HRO and its cause. Even with the use of a personalist frame, this information should reduce mobilization and support:

> **Research Hypothesis:** Exposure to discursive responses by governments will decrease support for the HRO's cause.

We think this hypothesis represents an important first step in understanding how government claims could influence support for an HRO's cause. Further work that examines differences in variation in claims is also necessary. For this first step, we collected original data using an online survey experiment that was administered from August 24, 2015 to August 31, 2015. Our sample approximates a U.S. nationally representative sample and consists of 405 participants.

Data and Method

Our online survey experiment uses the case of Nabeel Rajab. Nabeel Rajab is a prominent Bahraini human rights activist. He is the president of the Bahrain Center for Human Rights, founder of the Gulf Center for Human Rights, and Deputy Secretary General of the International Federation for Human Rights.[4] He is an opposition leader

and was among the three leaders of the Bahraini uprising as part of the wider Arab Spring (Leigh 2012). Mr. Rajab is a controversial figure (Kinninmont 2012) and has been labeled a terrorist by the government (Sandels 2010), but condemns violent protest (Leigh 2012). Most recently, he served a two-year prison sentence for charges relating to freedom of expression, and was released in May 2014. Mr. Rajab was then sentenced to six months in prison for tweeting, lost an appeal in 2015, and was later released, on health-related grounds, in July 2015. He was rearrested in June 2016, despite his deteriorating health.[5] On account of the remaining charges, which are all in response to tweets he posted, he has been sentenced to five years in prison.[6] The European Parliament, the UN Office of the High Commissioner for Human Rights, and the United States, among numerous other governmental and nongovernmental entities, have continuously called for his release; many entities have done so every time he was detained (Bahrain Center for Human Rights 2015a, 2015b).

The case of Mr. Rajab is appropriate for our experiment because he is (1) a prisoner of conscience (Amnesty International 2015); (2) a victim of human rights abuses, ranging from beatings by government agents to torture in prison (Agence France-Presse 2012); (3) an internationally recognized human rights activist who receives much support from both HROs and foreign governments when he is either imprisoned or charged; and (4) a target of relentless smearing campaigns by the authorities who have accused him of racism, sectarianism, violence, atheism, and terrorism (Kinninmont 2012; Sandels 2010). The experiment consists of a vignette and a set of questions. The vignette is a brief passage that describes the case of Mr. Rajab and refers to an Amnesty International call for the authorities to drop all charges against him. We referred to an Amnesty International call for action because of Amnesty's high brand recognition and because it has, in fact, sponsored urgent action calls to free and drop all charges against Mr. Rajab (Amnesty International 2015). The vignette has wording that is similar to news coverage of Amnesty International calls for action.

Participants were randomly assigned to one of two vignettes. Both vignettes use an identical personal frame with identical text, but differ in one attribute: the treatment vignette contains a typical government response to HRO shaming. It mentions that Mr. Rajab was labeled a terrorist by the Bahraini government. The control vignette does not. Our outcome measures aim to capture how upsetting participants find the case of Mr. Rajab, the extent to which they trust the information presented by the vignette, their willingness to sign a "call for action" petition, as well as their intent to donate to Amnesty International in support of this cause.

We collected data from 405 participants from a U.S. nationally representative sample in the summer of 2015.

We used the Qualtrics platform, through which participants were compensated for their time. Participants were paid seventy-five cents, resulting in an hourly wage of approximately six dollars, which is more than the median hourly reservation wage ($1.38 per hour) estimated by Horton and Chilton (2010). The median survey completion time was 7.4 minutes. The Qualtrics online platform belongs to a group of commercial platforms that generate online sampling frames and greatly facilitate conducting survey research. While it outsources the process of recruiting participants to other firms, Qualtrics can produce samples with specific democratic attributes that are often within a 10 percent range of their corresponding values in the U.S. population (Heen, Lieberman, and Miethe 2014). The platform is both efficient and relatively low cost when compared to traditional methods of gathering survey data.

Potential subjects were able to participate if they were older than eighteen years and were not familiar with the case of Mr. Rajab in advance. Participants were first asked a series of basic population questions, establishing age, income and education levels, race and ethnic identity, religious and political affiliation, and occupation.

A series of control questions followed. We asked participants whether they had ever served in the military, whether they had ever been deployed, and if yes, where and for how long. We asked the same about their family members and close friends. These questions aim to control for the possibility of combat-related effects. Recent research on political effects of combat at the level of the individual has produced conflicting findings. Combat experience has been linked not only to increased political mobilization and participation (Blattman 2009) and increased volunteerism (Nesbit and Reingold 2011), but also to a decrease in support for peaceful conflict resolution and to a lesser extent, a decrease in support for HROs (Grossman, Manekin, and Miodownik 2015). Basic information on military involvement will provide a modest check for possible combat effects.

Next, participants were asked whether they had ever participated in a protest, and if yes, whether it was violent. We also asked participants how often they use Twitter. Research in psychology shows that people tend to systematically favor and empathize with members of their own groups (ingroups) (Avenanti, Sirigu, and Aglioti 2010; Azevedo et al. 2013; Dasgupta 2004; Tajfel et al. 1971). As individuals require very little to form ingroups—random assignment of group identity based on T-shirt color will do (Wright et al. 1997)—it is reasonable to expect that participants who themselves engage in protest or use Twitter might more easily connect to Mr. Rajab than those who do neither. As members of the same ingroup—one that engages in protest or one that uses Twitter, or perhaps both—such participants are thus more likely to react

---

Amnesty International Urgent Action: Drop Charges Against Nabeel Rajab

Nabeel Rajab is a human rights activist and opposition leader from Bahrain. He became involved in human rights advocacy in early 1990s, and subsequently became one of the leading voices against human rights abuses in Bahrain. In Bahrain, however, he is seen as a controversial figure. He led several protests during the 2011 uprising, during which he clashed with the political authorities and security forces. In July 2012, he was arrested on charges of "publicly vilifying the people of al-Muharraq and questioning their patriotism with disgraceful expressions posted via social networking websites." During the trial Nabeel Rajab complained of physical and psychological torture in prison. He was acquitted of the charge, but remained in custody on three other charges, all related to illegal gathering and calling for unauthorized marches. He was released in May 2014, having served his full term of 2 years.

In April 2015, he was again arrested on the charges of "spreading false news." Later, two more charges were added, one of "insulting a statutory body" and the other of "disseminating false rumors in time of war." All three refer to Nabeel Rajab's tweets. Mr. Rajab faces up to 15 years of imprisonment under the charges relating to his tweets.

Amnesty International claims that Mr. Rajab has been unfairly punished for posting tweets, and that his right to free speech has been grossly violated. Amnesty International is calling for the Bahraini authorities to immediately drop all the charges against Mr. Rajab.

Figure 1. Control vignette.

strongly to the case of Mr. Rajab. In addition, participants from the United States might react particularly strongly as they can generally speak, protest, and/or tweet without fearing a years-long prison sentence.

Finally, we showed the participants the United Nations definition of torture and asked them whether they think that torture in prisons is acceptable. The purpose of this question is to establish an individual's baseline opinion on torture, which will aid in linking any difference between the treated and control groups to the "labeled terrorist" treatment. Suppose a participant who finds torture acceptable in certain situations receives the treatment vignette that includes the "labeled terrorist" statement, and that this participant also indicates not being upset by the case of Mr. Rajab. Without knowing that the participant finds torture acceptable in some situations, we might mistakenly attribute her reaction entirely to the treatment, while in reality her reaction might be due to her opinion on torture in general. In addition, establishing a baseline opinion on torture will help us determine the proportion of people who find that torture is never acceptable who are *also* willing to take action when presented with an opportunity to do so.

Participants were then randomly assigned to one of the two vignettes, which are fully presented in Figures 1 and 2. The vignettes describe Mr. Rajab as a human rights activist and opposition leader, mention his extensive involvement in protests, state that he had been imprisoned for social media activity and reportedly tortured in prison, and describe the new charges he is facing for tweeting. In writing the vignettes, we consulted source materials from Amnesty International, but wrote an original passage that used a personal frame, presented his current charges and most recent arrests, and, crucially, mentioned specific information on rights violations (torture) he allegedly suffered, which recent calls to action in the case of Mr. Rajab do not mention.

After the vignettes, we first posed an additional control question: we asked how knowledgeable participants felt about the case of Mr. Rajab (recall that we pre-screened potential participants and only included those who had no prior knowledge of Mr. Rajab). This question intends to control for the possibility of low levels of support for Mr. Rajab's that are not due to the "labeled terrorist" treatment but simply to the fact that a participant does not yet feel informed enough to be able to act. Answers to this question also have a practical implication for structuring calls for action; striking a balance between giving too little detail and constructing a passage that is too long might be crucial to mobilizing success.

Five outcome questions followed. The first aimed to establish the extent to which participants trusted the information presented. This measure serves both as a control and an outcome of interest. In its capacity as a control, the question helps determine whether an individual's lack of support for Mr. Rajab is due to the treatment or whether it is due to general mistrust of the HRO. As an outcome of interest, this question speaks to potential consequences of a damaging claim by a repressive government. While claims aimed at discrediting prisoners of conscience are perhaps primarily intended to erode support for that individual, they are just as often targeting the HRO advocating on behalf of that individual. This question therefore also aims to establish the extent to which an HRO might lose the trust of an audience based on a damning claim.

The second outcome question asked for a basic response to the case of Mr. Rajab. Participants were asked how upset they feel about Mr. Rajab being imprisoned for tweets that the authorities found offensive; possible answers were on a five-point scale ranging from "very upset" to "not at all upset." This question intends to measure (1) whether the vignettes were successful at spotlighting the rights abuse case (via the control), and (2) to

---

Amnesty International Urgent Action: Drop Charges Against Nabeel Rajab

Nabeel Rajab is a human rights activist and opposition leader from Bahrain. He became involved in human rights advocacy in early 1990s, and subsequently became one of the leading voices against human rights abuses in Bahrain. In Bahrain, however, he is seen as a controversial figure and has been declared a terrorist by the government. He led several protests during the 2011 uprising, during which he clashed with the political authorities and security forces. In July 2012, he was arrested on charges of "publicly vilifying the people of al-Muharraq and questioning their patriotism with disgraceful expressions posted via social networking websites." During the trial Nabeel Rajab complained of physical and psychological torture in prison. He was acquitted of the charge, but remained in custody on three other charges, all related to illegal gathering and calling for unauthorized marches. He was released in May 2014, having served his full term of 2 years.

In April 2015, he was again arrested on the charges of "spreading false news." Later, two more charges were added, one of "insulting a statutory body" and the other of "disseminating false rumors in time of war." All three refer to Nabeel Rajab's tweets. Mr. Rajab faces up to 15 years of imprisonment under the charges relating to his tweets.

Amnesty International claims that Mr. Rajab has been unfairly punished for posting tweets, and that his right to free speech has been grossly violated. Amnesty International is calling for the Bahraini authorities to immediately drop all the charges against Mr. Rajab.

Figure 2. Treatment vignette, with "labeled terrorist" statement.

what extent a damaging statement can reduce the intensity of a participant's response and so negatively affect the public perception of an action as a human rights violation (via the treatment). This outcome measure speaks to one of the primary aims in HRO mobilization efforts at the level of the individual: mobilizing consensus about the nature and severity of human rights violations.

The final three outcome measures aim to establish a participant's intent to act. These measures speak to another of the primary aims in HRO mobilization efforts: making individuals act, either to directly pressure perpetrators to stop rights violations or to help other actors in pressuring, shaming, or sanctioning the perpetrators. Our outcome measures vary in intensity of action. The first, least intense, asked participants how likely they were to support a campaign to drop the charges against Mr. Rajab, with possible answers on a five-point scale, ranging from "definitely would not support" to "definitely would support." This is an indirect measure of the intent to act.

The second outcome measure asked participants whether they would be willing to add their name to a petition calling for the Bahraini authorities to immediately drop all the charges against Mr. Rajab. This is a direct measure of the intent to act; participants had the option of answering either yes or no. These outcome questions are closely modeled after those in McEntire, Leiby, and Krain (2015).

We added a third, stricter outcome measure. Under this question, participants were invited to "TAKE ACTION NOW: Donate and help Amnesty International fight for Mr. Rajab!" and support the HRO fight for Mr. Rajab by donating. They had the option of clicking "No, thank you," which concluded their participation or "Yes, I want to take action!" which took them to an outside link to Amnesty International donations page (in addition to concluding their participation). As we do not track actual donations and only capture whether participants want to take action, we only capture their intent. Nonetheless, this is a fairly sensitive measure as the step between willingness to add your name to a petition at no material cost and the willingness to go through the motions of making a donation and the cost of the donation itself is a significant one. With regard to our treatment, it is possible that audiences who are otherwise willing to financially support a cause might be dissuaded from doing so if the integrity of the political prisoner is questioned.

Results

In this section, we present results from our analysis evaluating the effect of government counterclaims leveled at prisoners of conscience on HRO mobilizing success. Our findings suggest that a terrorist label does decrease support for a cause, but not across the board. Participants who were assigned the "labeled terrorist" statement were significantly less likely to sign a petition asking for the Bahraini authorities to drop all charges against Mr. Rajab and indicated less support for such a campaign in general; they were also less likely to feel upset by the case. We find that the treatment does not negatively affect participants' trust in the information presented or their willingness to donate money to Amnesty International in support of the cause.

The results of logit and ordered logit regressions are presented in Table 1. The control variables included in this analysis are limited to the levels of personal engagement in protest (0—*never*, 4—*more than ten*), the frequency of Twitter use (0—*never*, 7—*every hour*), opinion regarding acceptability of torture in prisons (0—*never acceptable*, 5—*always acceptable*), political affiliation (0—*very conservative*, 7—*very liberal*), income level, gender, age, and level of education. We present the results of regressions that use each of the five outcome measures as a dependent variable: trust in information provided (0—*no trust*, 4—*complete trust*), feelings toward Mr.

Table 1. Effect of Terrorist Labeling on Support for Human Rights Organization Campaign.

Variables	(1) Trust information	(2) Feelings about case	(3) Campaign support	(4) Sign petition	(5) Donate
Terrorist Labeling Treatment	−0.01157 (0.189)	−0.40302** (0.183)	−0.37946** (0.181)	−0.70246*** (0.219)	−0.14467 (0.294)
Protest	0.34066*** (0.116)	0.25462** (0.121)	0.40069*** (0.116)	0.68692*** (0.153)	0.58406*** (0.147)
Twitter	0.16388*** (0.045)	0.04882 (0.051)	0.22588*** (0.047)	0.20068*** (0.056)	0.20391*** (0.066)
Torture Acceptability	−0.19343** (0.079)	−0.39054*** (0.083)	−0.24144*** (0.075)	−0.23886*** (0.084)	−0.17376 (0.108)
Political Affiliation	−0.00475 (0.045)	0.08908** (0.044)	0.01778 (0.042)	0.01088 (0.051)	0.03702 (0.063)
Income	−0.00861 (0.028)	0.02392 (0.025)	0.01039 (0.024)	0.00307 (0.030)	−0.00978 (0.037)
Gender	−0.53789*** (0.204)	0.04635 (0.194)	−0.31026 (0.194)	−0.25273 (0.226)	−0.51317* (0.303)
Age	0.01875 (0.033)	0.02490 (0.032)	−0.01444 (0.032)	0.01828 (0.037)	−0.04136 (0.049)
Education	0.09436* (0.048)	0.00239 (0.049)	−0.00817 (0.043)	−0.05487 (0.058)	−0.01508 (0.076)
Constant Cut 1	−1.75880*** (0.674)	−3.21099*** (0.762)	−2.59265*** (0.785)		
Constant Cut 2	−0.77695 (0.680)	−1.80224** (0.734)	−1.49920** (0.750)		
Constant Cut 3	1.38872** (0.682)	−0.09242 (0.740)	0.37265 (0.745)		
Constant Cut 4	3.24708*** (0.730)	1.33676* (0.743)	1.82379** (0.766)		
Constant				−0.21942 (0.845)	−1.68255* (0.989)
Observations	405	405	405	405	405

Robust standard errors in parentheses.
*$p < .1$. **$p < .05$. ***$p < .01$.

Rajab's case (0—*not at all upset*, 4—*very upset*), willingness to support a campaign to drop all charges (0—*no support*, 4—*full support*), willingness to sign a petition to drop all charges (yes/no), and willingness to donate to Amnesty International in support of this cause (yes/no).

First, as shown in Table 1, column 2, participants who received the "labeled terrorist" treatment were significantly less likely to be upset by the case of Mr. Rajab. For example, participants in the treated group were on average 7.5 percent less likely to feel "very upset" about the case. A full table of the marginal effects of the treatment across responses is provided in Table 2 and illustrated in Figure 3, which provides the average marginal effects of discrete changes in all the covariates. This finding suggests that damaging statements may negatively influence an HRO's ability to mobilize consensus about the nature and severity of a violation (Table 2).

Second, as shown in column 3 of Table 1, the treatment also appears to negatively affect participants'

willingness to support a campaign to drop all charges against Mr. Rajab. Reading one damaging sentence, for example, results in a 4.6 percent decrease in full support for such a campaign. This suggests that a damaging statement can affect an HRO's ability to mobilize at the very basic level, which aims to secure general support for a campaign but does not yet ask people to act. Table 3 provides the marginal effects of the treatment on all campaign support outcomes. Figure 4 shows the average marginal effects for all covariates.

Third, as shown in column 4 of Table 1, just as a "terrorist label" reduces support for a campaign to drop all charges against Mr. Rajab, it reduces the willingness of participants to sign a petition to that effect. Having seen such a label, participants in the treated group are 14.7 percent less likely to sign a petition for Mr. Rajab. These marginal effects are provided in Table 4 and Figure 5. The damage that a "terrorist label" can inflict is therefore palpable and occurs quite early in the process of mobilization. It appears to
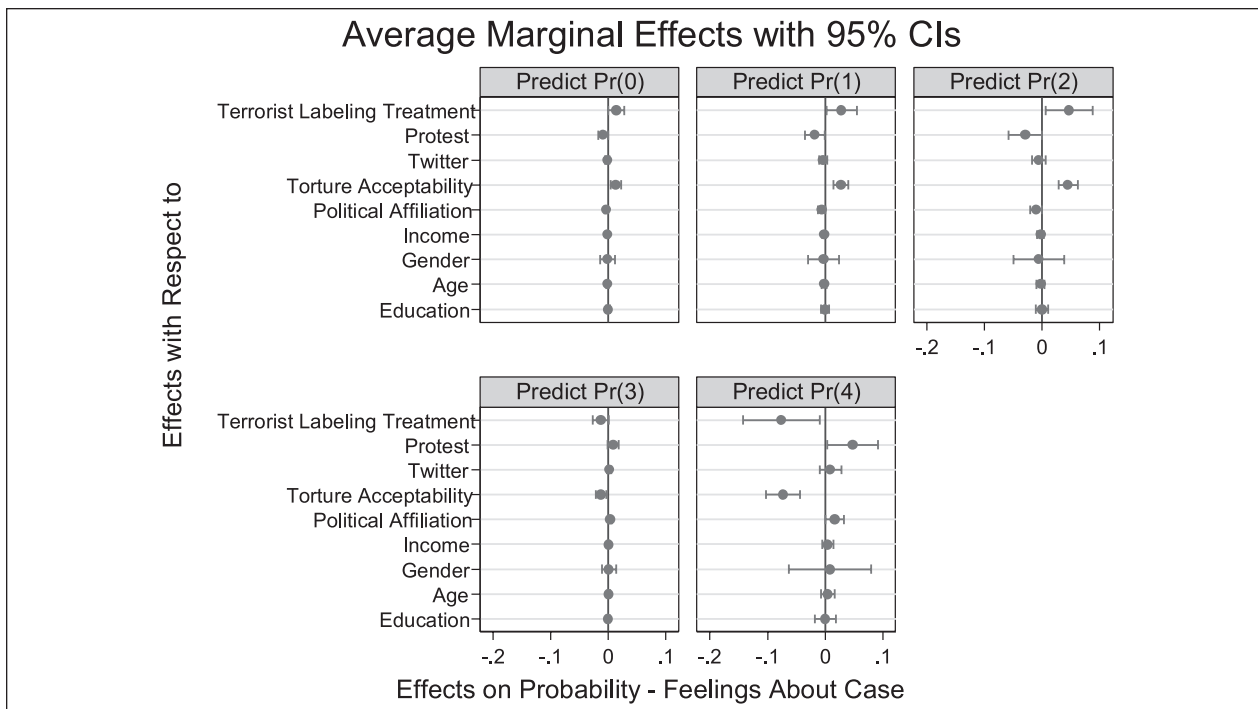
Figure 3. Average marginal effects, feelings toward Mr. Rajab's case.
CI = confidence interval.

Table 2. Marginal Effect of Terrorist Labeling Treatment on Feelings toward Mr. Rajab's Case.

	Feel
Predict Pr(0—*not upset at all*)	0.013
	(0.007)
Predict Pr(1)	0.028
	(0.013)*
Predict Pr(2)	0.047
	(0.021)*
Predict Pr(3)	−0.013
	(0.007)
Predict Pr(4—*very upset*)	−0.075
	(0.034)*
N	405

*p < .1. **p < .05.

Table 3. Marginal Effect of Terrorist Labeling Treatment on Campaign Support.

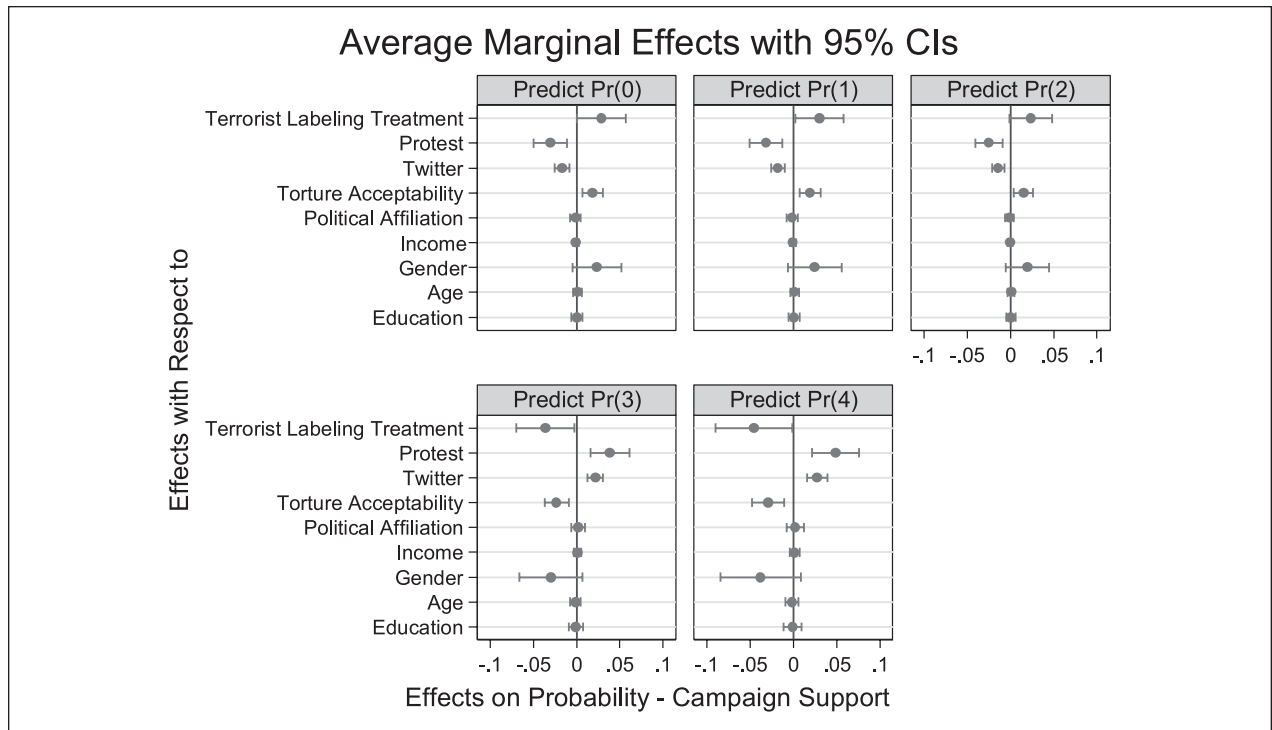|	Campaign support
Predict Pr(0—*no support*)	0.029
	(0.014)*
Predict Pr(1)	0.030
	(0.014)*
Predict Pr(2)	0.024
	(0.013)
Predict Pr(3)	−0.036
	(0.017)*
Predict Pr(4—*full support*)	−0.046
	(0.023)*
N	405

*p < .1. **p < .05.

dissuade audiences who might otherwise participate in a campaign from supporting it, and then also from participating in it with their signature.

The treatment does not, however, appear to have a negative effect on the willingness to donate to Amnesty International in support of Mr. Rajab, as shown in Table 1, column 5. This finding is in line with other research that finds that personalist frames, in general, do not mobilize donations in the same way as they mobilize individuals to sign petitions or alter opinions (Baron and Szymanska 2011; Cryder and Loewenstein 2011; McEntire, Leiby,

and Krain 2015). As the overall number of participants who indicate willingness to donate is quite low (*N* = 63), however, the lack of an effect may also be due to insufficient variation.

Finally, the findings suggest that a damaging statement about a prisoner of conscience does not reduce individual trust in the information provided by an HRO, as shown in column 1 of Table 1. An HRO advocating for such an individual, then, may not itself be vulnerable to shaming attempts by repressive governments. Future

Figure 4. Average marginal effects, campaign support.
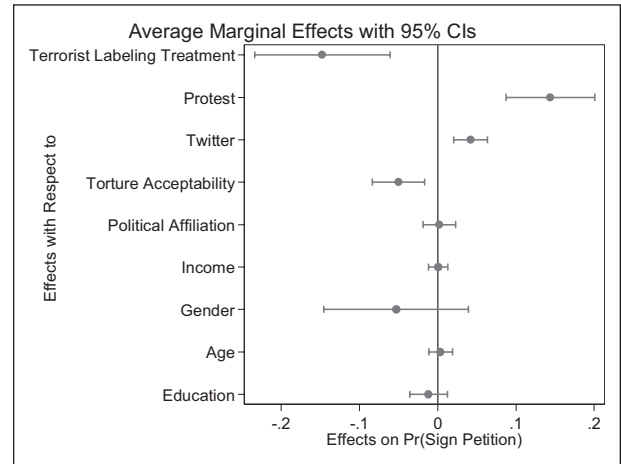CI = confidence interval.

Table 4. Marginal Effect of Terrorist Labeling Treatment on Willingness to Sign Petition for Mr. Rajab.

	Sign petition
Terrorist Labeling Treatment	−0.147
	(0.044)**
N	405

*p < .1. **p < .05.

surveys could draw on this finding and examine whether respondents' trust in specific parts of the vignette or the organization are influenced by the treatment.

While the relationship between a terrorist label and lessened support for the HRO campaign might not seem surprising, our findings are remarkable for two reasons. First, the treatment is subtle. The rather lengthy ~250 word vignettes differ by nine words: "and has been declared a terrorist by the government." Second, the treatment statement is not necessarily credible. As a repressor, the Bahraini government might not be perceived as a credible source in this matter, particularly as the act of shaming a prominent human rights activist does not run against type. Dismissing the claims of an HRO or countering them with a damaging counter-accusation is precisely what one might expect from a repressive government. Similarly, when a conservative media

Figure 5. Average marginal effects, sign petition.
CI = confidence interval.

outlet praises a Republican president, it is not running against type and thus faces credibility challenges; were such an outlet to run against type by criticizing a Republican president, however, its statements would be more influential (Baum and Groeling 2009). Yet, even as the Bahraini government's labeling of Mr. Rajab as a terrorist is nothing out of the ordinary, the label still results in a sizeable effect.

Among the control covariates, three stand out as particularly significant. Personal participation in protest is statistically significant across the board. Participants who participated in protests, marches, or rallies themselves are more likely to feel upset by the case of Mr. Rajab, more likely to sign a petition and donate, more likely to support the campaign in general, and more likely to trust the information presented. Personal use of Twitter is significant in four out of five cases; participants who use Twitter often are more likely to sign a petition and donate, more likely to support the campaign in general, more likely to trust the information presented, but no more likely to feel upset by the case than are those who use Twitter less often or not at all. Finally, views on the acceptability of torture in prisons are related, in the expected direction, to all outcomes but one: participants who find the use of torture unacceptable are no more and no less likely to donate to the campaign than those who find it acceptable.

Also significant, but not consistently, are covariates on gender, education, political affiliation, and race. Women are more likely to trust the information presented, as are people with higher levels of education. Women are also more likely to donate to the campaign. Participants who identify as liberal are more likely to be upset about the case than are participants who identify as conservative.

Together, these findings suggest that repressive governments can damage the effectiveness of HRO action by attempting to discredit a prisoner of conscience. They further suggest that the point of vulnerability is not centered around donations, but earlier in the process of mobilization. Potential supporters of a campaign are likely to be dissuaded from *merely supporting a campaign* and then *putting their name on a list* because of a negative label. The power to turn away potential advocates at an early stage can considerably impact HRO advocacy and the success of the advocacy network, even if the HRO's own image is not threatened by a damaging accusation.

Conclusion

Human rights campaigns are very political phenomena, with much back and forth between HROs and governments. HROs report facts about government abuses as they see it; governments typically strike back with attacks against the HROs and the individuals they represent. Often, these attacks include labeling the abused as a terrorist, even without any evidence that that is actually the case. Many times, both the HRO's report and the government's response are presented in the international press.

In this project, we use an experimental research design to examine whether a discursive response by an accused government, this time labeling a prisoner of conscience a terrorist, reduces individual-level support to the HRO's campaign. Drawing on McEntire, Leiby, and Krain

(2015), we focus on a personalist frame, where individuals are told in a passionate way about the abuse the prisoner of conscience suffers. They find that this information alone should spur individuals into action.

As we argue and find, however, HROs aren't the only actors passionately making their case. When governments respond to HRO claims by labeling the abused as a terrorist, individuals are much less likely to be spurred into certain types of action, like signing a petition, and feel differently about the specific case. Interestingly, we do not find that trust in the information provided by an HRO is harmed by being associated with someone labeled a terrorist.

These results are important to the larger academic and activist communities in a number of regards. First, our findings speak to the larger cross-disciplinary literatures on empathy (Azevedo et al. 2013). We find that even impassioned personalist accounts of abuse can be derailed by references to terrorism, especially among certain groups of individuals that see themselves as disconnected from the prisoner of conscience.

Second, our findings speak to how crucial the discursive process between HROs and governments can be and how difficult mobilization of the transnational advocacy network really is. If individual-level support for demanded human rights changes decreases as a result of unsubstantiated claims by the targeted government, moving through Risse, Ropp, and Sikkink's (1999) phases of human rights improvements is extremely difficult. If individual support is less likely for abuse victims that are labeled terrorists, this could create a situation where HROs shift their focus away from those that have been branded terrorists, whether that labeling is correct or not. This finding thus provides a cautionary tale for HRO activists about potential structural limits to their work.

Furthermore, although our work focused on human rights, it is likely that the processes we identified also occur in other advocacy areas, like in advocacy related to the environment. Similar to labeling a victim a terrorist, governments often label environmental causes as harmful to economic development. Future work could examine whether these counterclaims similarly limit mobilization.

The findings also highlight the role of the media in the mobilization process. Governments may make counterclaims to HRO reports of abuses; as discussed above, some of these counterclaims may be false. However, in an effort to appear unbiased, the media may report the government's claims alongside the HRO information. In doing so, the media could be ultimately helping to limit support for the abused. Scholarship in this area could benefit with further theorizing about the important role of the media in disseminating and possibly altering activist messages.

Finally, this research, like Ausderan (2014) and McEntire, Leiby, and Krain (2015), stresses the necessity

of unpacking the black box of transnational advocacy networks and looking at individuals and their likelihood of support for human rights causes. Early research on transnational advocacy networks was essential for moving the field of international relations beyond the traditional focus on states and state-to-state relations and taking seriously the role played by non-state organizations, like HROs (Keck and Sikkink 1998; Risse, Ropp, and Sikkink 1999). Future work examining how individuals respond to HRO messaging and the ways in which individuals mobilize into existing and new organizations has much potential for the field of human rights and speaks to perennial questions in contentious politics. When are individuals likely to support nongovernmental causes? When are they prepared to act on that support? What can governments do to respond? This project looks at these questions with a focus on the discursive tools governments have to limit critiques against them and thwart mobilization that could rein in their repressive practices. Unfortunately, for HROs, human rights advocates, and victims of abuse, governmental efforts to deny the charges against them are at least somewhat effective at limiting the potential advocacy network.

Declaration of Conflicting Interests

Funding

Notes

1. Notable exceptions to this include Risse, Ropp, and Sikkink (1999) and Vennesson and Rajkovic (2012). As Vennesson and Rajkovic (2012) state, our understanding of the "processes of strategic interaction between transnational advocates and their targeted state actors" is very limited (410).
2. http://www.irishtimes.com/news/politics/egypt-unable-to-aid-halawa-until-terrorism-case-ends-1.2303963, https://www.bbc.com/news/world-europe-41310237
3. http://www.iwmf.org/reeyot-alemu-2012-courage-in-journalism-award/
4. http://www.bahrainrights.org/en/human-rights-defenders-targeted-bahrain-2

5. http://www.bahrainrights.org/en/updates-arrest-and-detention-bchrs-president-nabeel-rajab
6. http://www.bahrainrights.org/en/node/7571, https://www.frontlinedefenders.org/en/case/case-history-nabeel-rajab

ORCID iD

Amanda Murdie https://orcid.org/0000-0002-6459-0594

Supplemental Material

Replication materials for this article are available at amandamurdie.org.

References

Agence France-Presse. 2012. "Bahrain Activist Acquitted but Remains in Jail." August 23. http://sputniknews.com/voiceofrussia/2012_08_23/Bahrain-activist-acquitted-but-remains-in-jail/ (accessed August 27, 2015).

Amnesty International. 2015. "Urgent Action: Nabeel Rajab Loses Appeal against Sentence." http://www.amnestyusa.org/sites/default/files/uaa24914_6.pdf (accessed August 27, 2015).

Ausderan, Jacob. 2014. "How Naming and Shaming Affects Human Rights Perceptions in the Shamed Country." *Journal of Peace Research* 51 (1): 81–95.

Avenanti, Alessio, Angela Sirigu, and Salvatore M. Aglioti. 2010. "Racial bias reduces empathic sensorimotor resonance with other-race pain." *Current Biology 20*(11): 1018–1022.

Azevedo, Ruben T., Emiliano Macaluso, Alessio Avenanti, Valerio Santangelo, Valentina Cazzato, and Salvatore Maria Aglioti. 2013. "Their Pain is not Our Pain: Brain and Autonomic Correlates of Emphatic Resonance with the Pain of Same and Different Race Individuals." *Human Brain Mapping* 34:3168–81.

Bahrain Center for Human Rights. 2015a. "Bahrain: Release of Leading Human Rights Defender Nabeel Rajab Must Be Followed with More Steps for Consistent Protection of Human Rights." http://www.bahrainrights.org/en/node/7571 (accessed August 27, 2015).

Bahrain Center for Human Rights. 2015b. "Human Rights Defenders Targeted in Bahrain." http://www.bahrainrights.org/en/human-rights-defenders-targeted-bahrain-2 (accessed August 27, 2015).

Baron, Jonathan, and Ewa Szymanska. 2011. "Heuristics and Biases in Charity." In *The Science of Giving: Experimental Approaches to the Study of Charity*, edited by Daniel M. Oppenheiemer and Christopher Y. Olivola, 215–35. New York, NY: Psychology Press.

Barry, Colin M., K. Chad Clay, and Michael E. Flynn. 2013. "Avoiding the Spotlight: Human Rights Shaming and Foreign Direct Investment." *International Studies Quarterly* 57 (3): 532–44.

Baum, Matthew A., and Tim J. Groeling. 2009. *War stories: The Causes and Consequences of Public Views of War*. Princeton, NJ: Princeton University Press.

BBC News. 2005. "Bush says Amnesty report 'absurd'." http://news.bbc.co.uk/go/pr/fr/-/2/hi/americas/4598109.stm (accessed June 25, 2019).

Blattman, Christopher. 2009. "From Violence to Voting: War and Political Participation in Uganda." *American Political Science Review* 103 (2): 231–47.

Bob, Clifford. 2005. *The Marketing of Rebellion: Insurgents, Media, and International Activism*. Cambridge: Cambridge University Press.

Bouvard, Marguerite Guzman. 1994. *Revolutionizing Motherhood: The Mothers of the Plaza de Mayo*. Lanham: Rowman & Littlefield.

Brysk, Alison. 2014. *Speaking Rights to Power: Constructing Political Will*. Oxford: Oxford University Press.

Carothers, Thomas. 2016. "Closing Space for International Democracy and Human Rights Support." *Journal of Human Rights Practice* 8 (3): 358–77.

Carothers, Thomas, and Saskia Brechenmacher. 2014. *Closing Space: Democracy and Human Rights Support under Fire*. Washington, DC: Carnegie Endowment for International Peace.

Carpenter, R. Charli. 2007. "Setting the Advocacy Agenda: Theorizing Issue Emergence and Nonemergence in Transnational Advocacy Networks." *International Studies Quarterly* 51 (1): 99–120.

Cryder, Cynthia E., and George Loewenstein. 2011. "The Critical Link between Tangibility and Generosity." In *The Science of Giving: Experimental Approaches to the Study of Charity*, edited by Daniel M. Oppenheiemer and Christopher Y. Olivola, 237–52. New York, NY: Psychology Press.

Dasgupta, Nilanjana. 2004. "Implicit Ingroup Favoritism, Outgroup Favoritism, and Their Behavioral Manifestations." *Social Justice Research* 17 (2): 143–69.

Davis, David R., Amanda Murdie, and Coty Garnett Steinmetz. 2012. "Makers and Shapers: Human Rights INGOs and Public Opinion." *Human Rights Quarterly* 34 (1): 199–224.

DeMeritt, Jacqueline H. R. 2012. "International Organizations and Government Killing: Does Naming and Shaming Save Lives?" *International Interactions* 38 (5): 597–621.

Grossman, Guy, Devorah Manekin, and Dan Miodownik. 2015. "The Political Legacies of Combat: Attitudes toward War and Peace among Israeli Ex-Combatants." *International Organization* 69:981–1009.

Hafner-Burton, Emilie, and James Ron. 2013. "The Latin Bias: Regions, the Anglo-American Media, and Human Rights." *International Studies Quarterly* 57 (3): 474–91.

Heen, Miliaikeala S. J., Joel D. Lieberman, and Terance D. Miethe. 2014. "A Comparison of Different Online Sampling Approaches for Generating National Samples." *University of Nevada, Las Vegas Center for Crime and Justice Policy: Research in Brief*. www.unlv.edu/ccjp (accessed August 27, 2015).

Hendrix, Cullen S., and Wendy H. Wong. 2013. "When Is the Pen Truly Mighty? Regime Type and the Efficacy of Naming and Shaming in Curbing Human Rights Abuses." *British Journal of Political Science* 43 (3): 651–72.

Hill, Daniel W., Will H. Moore, and Bumba Mukherjee. 2013. "Information Politics versus Organizational Incentives: When Are Amnesty International's 'Naming and Shaming' Reports Biased?" *International Studies Quarterly* 57 (2): 219–32.

Horton, John J., and Lydia B. Chilton. 2010. "The Labor Economics of Paid Crowdsourcing." EC'10 proceedings of the 11th ACM Conference on Electronic Commerce, Association for Computing Machinery, New York, 209–18.

Hortsch, Diana. 2010. "The Paradox of Partnership: Amnesty International, Responsible Advocacy, and NGO Accountability." *Columbia Human Rights Law Review* 42:119–50.

Keck, Margaret E., and Kathryn Sikkink. 1998. *Activists beyond Borders: Advocacy Networks in International Politics*. Ithaca, NY: Cornell University Press.

Kinninmont, Jane. 2012. "Bahrain's still Stuck." *Foreign Policy*, August 15. http://foreignpolicy.com/2012/08/15/bahrains-still-stuck/ (accessed August 27, 2015).

Krain, Matthew. 2012. "J'accuse! Does Naming and Shaming Perpetrators Reduce the Severity of Genocides or Politicides?" *International Studies Quarterly* 56 (3): 574–89.

Leigh, Karen. 2012. "The Agony of Nabeel Rajab." *The Atlantic*, February 13. http://www.theatlantic.com/international/archive/2012/02/the-agony-of-nabeel-rajab/252493/ (accessed August 27, 2015).

McEntire, Kyla Jo, Michele Leiby, and Matthew Krain. 2015a. "How to Ask People for Change: Examining Peoples' Willingness to Donate to Human Rights Campaigns." In *The Social Practice of Human Rights*, edited by Joel Pruce, 43–62. New York, NY: Palgrave Macmillan.

McEntire, Kyla Jo, Michele Leiby, and Matthew Krain. 2015b. "Human Rights Organizations as Agents of Change: An Experimental Examination of Framing and Micromobilization." *American Political Science Review* 109 (3): 407–26.

McEntire, Kyla Jo, Michele Leiby, and Matthew Krain. 2017. "How Combining Framing Strategies Affects Human Rights Micromobilization." *Research & Politics* 4 (2): 1–11.

Murdie, Amanda, and David R. Davis. 2012. "Shaming and Blaming: Using Events Data to Assess the Impact of Human Rights INGOs." *International Studies Quarterly* 56 (1): 1–16.

Murdie, Amanda, and Dursun Peksen. 2013. "The Impact of Human Rights INGO Activities on Economic Sanctions." *The Review of International Organizations* 8 (1): 33–53.

Murdie, Amanda, and Dursun Peksen. 2014. "The Impact of Human Rights INGO Shaming on Humanitarian Interventions." *The Journal of Politics* 76 (1): 215–28.

Nesbit, Rebecca, and David A. Reingold. 2011. "Soldiers to Citizens: The Link between Military Service and Volunteering." *Public Administration Review* 71:67–76.

NPR. "Is Amnsety International Supporting a Jihadist?" https://www.npr.org/templates/story/story.php?storyId=124156482 (accessed June 25, 2019).

Peksen, Dursun. 2009. "Better or Worse? The Effect of Economic Sanctions on Human Rights." *Journal of Peace Research* 46 (1): 59–77.

Peksen, Dursun. 2012. "Does Foreign Military Intervention Help Human Rights?" *Political Research Quarterly* 65 (3): 558–71.

Poe, Steven C. 1990. "Human Rights and US Foreign Aid: A Review of Quantitative Studies and Suggestions for Future Research." *Human Rights Quarterly* 12:499–512.

Richards, David L., Ronald D. Gelleny, and David H. Sacko. 2001. "Money with a Mean Streak? Foreign Economic Penetration and Government Respect for Human Rights in Developing Countries." *International Studies Quarterly* 45 (2): 219–39.

Risse, Thomas, Steven C. Ropp, and Kathryn Sikkink. 1999. *The Power of Human Rights: International Norms and Domestic Change*. Cambridge: Cambridge University Press.

Risse, Thomas, Steven C. Ropp, and Kathryn Sikkink. 2013. *The Persistent Power of Human Rights: From Commitment to Compliance*. Cambridge: Cambridge University Press.

Risse, Thomas, and Kathryn Sikkink. 1999. "The Socialization of International Human Rights Norms into Domestic Practices: Introduction." In *The Power of Human Rights: International Norms and Domestic Change*, edited by Thomas Risse and Steven C. Ropp,nd Kathryn Sikkink, 1–38. Cambridge: Cambridge University Press.

Ron, James, Howard Ramos, and Kathleen Rodgers. 2005. "Transnational Information Politics: NGO Human Rights Reporting, 1986–2000." *International Studies Quarterly* 49 (3): 557–88.

Sandels, Alexandra. 2010. "Bahrain Arrests Dozens of Shiite Muslims Accused of Plotting Coup." *The Los Angeles Times*, September 7. http://articles.latimes.com/print/2010/sep/07/world/la-fg-bahrain-crackdown-20100907 (accessed August 27, 2015).

Tajfel, Henri, Michael G. Billig, Robert P. Bundy, and Claude Flament. 1971. "Social Categorization and Intergroup Behaviour." *European Journal of Social Psychology* 1 (2): 149–78.

Vennesson, Pascal, and Nikolas M. Rajkovic. 2012. "The Transnational Politics of Warfare Accountability: Human Rights Watch versus the Israel Defense Forces." *International Relations* 26 (4): 409–29.

Weinraub, Bernard. 1982. "U.S. Considers Guatemala Arms Aid." *The New York Times*, November 24. http://www.nytimes.com/1982/11/24/world/us-considers-guatemala-arms-aid.html (accessed August 27, 2015).

Wilson, Elizabeth A. 2016. "Restrictive National Laws Affecting Human Rights Civil Society Organizations: A Legal Analysis." *Journal of Human Rights Practice* 8 (3): 329–57.

Wright, Stephen C., Arthur Aron, Tracy McLaughlin-Volpe, and Stacy A. Ropp. 1997. "The Extended Contact Effect: Knowledge of Cross-Group Friendships and Prejudice." *Journal of Personality and Social Psychology* 73 (1): 73–90.