International Diffusion of Development Knowledge:

A Field Experiment on Officials' Responses to Impact Evidence in Peru

Darren Hawkins

Aubriana Wolferts

Daniel L. Nielson

24 August 2018

Social scientists have chronicled global improvements across a wide array of development indictors including education, health, electrification, sanitation, roads, purchasing power, and peace-building, among many other categories (Kenny 2011, Deaton 2013, Pinker 2018, Rosling 2018). In explaining the global decrease in extreme poverty, scholarly consensus is coalescing around the application of knowledge and technologies that improve social welfare at diminishing cost. Yet many questions remain about the mechanisms behind the diffusion of these lifesaving and life-improving ideas and technologies. What motivates governments to adopt new policies and practices from other locales in ways that improve the lives and livelihoods of their citizens?

Policy diffusion takes place when government policy choices are shaped by the actions of other governments (Graham, Shipan and Volden 2013; Simmons, Dobbin and Garrett 2006; Maggetti and Gilardi 2016). Diffusion may occur between local or regional governments inside a state or between states. Aligned with the development policies above, scholars have identified a very wide range of policies that have diffused internationally, including tax policies, independent regulatory agencies, good governance, environmental policy, bilateral treaties, welfare systems, and many others (Maggetti and Gildardi 2016, 88). They have also noted four main mechanisms by which diffusion occurs: learning, competition, coercion and socialization (Graham, Shipan and Volden 2013, 690).

While diffusion is well-documented, the underlying processes remain somewhat more obscure (Graham, Shipan and Volden 2013, 697). Many diffusion studies focus on particular policies and actors rather than on the generalizable factors that might promote diffusion. As Graham, Shipan and Volden put it (2012, 697), "Future work needs to discern systematic patterns in the conditional nature of diffusion." Moreover, much of the diffusion literature has been focused on an aggregate governmental level of analysis (Butler, Volden, Dynes and Shor 2017). All policy diffusion involves individual people making decisions, often small and limited in

scope, that contribute to these aggregate patterns. We seek to understand these micro-foundations.

Diffusion processes require moments of inception, times when policymakers first demonstrate interest in understanding how others have acted. Insights into those moments of inception can shed light on policy diffusion processes because they help scholars understand why policymakers might first demonstrate any interest at all in diffusion from the very beginning. In all cases except the most blatant coercion, policymakers must make choices about whether to pursue information about the actions of others. We seek a better understanding of why policymakers might choose to learn more about others' policy ideas and outcomes.

We designed an experiment that presented policymakers in Peru with the opportunity to make such choices. We approached 2,739 policymakers in Peru with a promise to demonstrate a website providing information about academic studies of development-policy programs around the world. We randomly varied the nationality and gender of the messenger and the social proof of whether other policymakers were also interested in seeing our demonstration. We observed policymaker responses in inviting meetings with us and in actually holding those meetings. While messenger characteristics and social proof are not the only factors that might induce policymaker interest in starting the diffusion process, they are factors that have proven important in shaping many related types of behavior.

In our experiment, we hired student researchers to contact policymakers through emails and phone calls. We expected U.S. nationals to be more credible than locals due to their association with a more developed country with higher intellectual prestige and a reputation for high-quality science and policy. We expected government officials to respond better to invitations from men than from women, in part because social norms about credibility in employment are often gendered. Men are typically viewed as more credible in areas of science and academics, which were the issue areas at hand. We operationalized social proof by telling policymakers that

others in their ministry or government had already accepted an invitation to host a briefing. Those statements were always true, as we carefully tracked who had already responded to our invitations.

We find that messenger nationality had very large effects on policymaker response, with U.S. nationals receiving far more favorable responses compared to local policymakers. Gender, however, worked differently than we expected, as policymakers were less likely to respond negatively to invitations from women than men but otherwise displayed no significant difference across messenger gender in positive responses, briefings actually held, or number of officials trained. Also contrary to our expectations and to an enormous literature showing strong effects in a large array of issue areas, we found null results for our social proof treatments: those prompted that ministry peers had already accepted invitations for briefings were no more likely to accept invitations than control subjects. We suspect that social proof may work differently in a government workplace environment and perhaps in diffusion processes generally.

Empirical Context

Policy impact-evaluation studies provide an important medium for policy diffusion. In recent years, policy makers around the world have strongly promoted the use of such studies as an important tool for getting the policy right. Rigorous evaluation through randomized control trials (RCTs) has become the gold standard for assessing the effectiveness of development programs. Major donor organizations and developing countries have widely adopted the practice. Scholars and their governmental counterparts have produced thousands of these evaluations. In our surveys of policymakers conducted in conjunction with this research, policymaker interest in academic studies as a source of information ranked first among possible information sources.

At the same time, most research has shown that evidence-based policy has limited

influence on policymakers (Newman, Cherney and Head 2017; Cairney, Oliver and Wellstead 2016; Head 2015, 2016). The academic nature of these evaluations makes them difficult for policy makers to understand and to learn from. Rigorous impact evaluations are typically written for academic audiences, using arcane language and methods, difficult mathematical models, and insider jargon. Analysts differ greatly from each other in the way they present the information, so understanding the studies requires a great deal of time, effort, experience and sophistication, even for those with appropriate training. Moreover, the studies are often in English, a language of which a large number of government officials do not have deep mastery.

We sought to minimize these barriers by creating a free, easy-to-use website that summarizes the key findings of impact evaluations and presents them in officials' native language and in accessible graphical formats. Normatively, our research is driven by the view that if policymakers could easily access and understand the thousands of RCTs on development policy, they could use those studies as a reference while doing their work. This means they would be more likely to design and implement policies that are effective at promoting economic growth and public welfare on a large scale. To be sure, some scholarly organizations producing RCTs engage in policy outreach, which is growing. Still, more could be done. Most of these efforts, for example, do not present graphical results of the findings or translate summaries into native languages.

Our website, impactevidence.byu.edu, summarizes the key findings of more than 400 RCTs produced by some of the most well-known development-oriented research centers in the United States. It uses easy-to-understand graphs and favors those graphs over words in presenting the main results from each study. Our website contains information from RCTs across the world in a variety of fields including education, health, agriculture, finance, governance, and business. Policymakers in all sectors can apply results from these field

experiments to improve their constituents' lives.

Unfortunately, we know little about what motivates individual development policy makers to seek new information. In other words, questions persist about how this type of diffusion process starts. Most of the existing research focuses on systemic-level factors like bureaucratic standards and the relationship between public agencies and information providers (Head 2016). While those factors are undoubtedly important, it is also essential to understand what might motivate particular individuals to engage in seeking more information.

Our project also provides an important dimension of realism. In their daily tasks, government officials come across immense amounts of information and must process that information, sorting it based on utility. The studies on our website represent one type of information that government officials could encounter. We dispatched teams of researchers in country to promote our informational website with those officials. We thus engaged in the sorts of tasks that policy advocates and information providers typically carry out as they seek to influence policy processes.

We initially selected Peru, Tanzania and India as sites for our study. Unfortunately, due to general unavailability of email addresses for officials in Tanzania and to officials' almost total unresponsiveness to our email messages in India, we were compelled to exclude those two countries from the present analysis and to concentrate on the results in Peru.[1] In Peru, email addresses were widely available for government officials and response rates were good, so the Peru site offers a relatively clean test of the study's core hypotheses operationalized in the interventions. Hence, we focus only on Peru in this paper.

[1] We persisted to contact officials in India and Tanzania through phone- and in-person communication, where enumerators' local language ability critically confounded the key nationality intervention. For Peru, we translated the website into Spanish and hired only Spanish-speaking students as messengers, which largely eliminated the language confound encountered in the two other countries. We selected Tanzania and India in part because English is an official language in both countries. However, we over-estimated Indian and Tanzanian officials' English-language abilities, which required contact in the local language.

Theoretical Context

What increases government officials' interest in learning more about development policies around the world? More generally, what influences the inception of diffusion? We focus on two factors: credibility and social proof. Scholars have produced copious literatures on how these factors influence individual decision-making, though the literature is less extensive in political environments and almost entirely lacking with respect to diffusion decisions.

Any new information faces a substantial credibility problem, perhaps especially in a political or policy-making context. Why should government officials be motivated to examine the policy decisions of others? Busy policy makers with serious time and resource constraints must make quick decisions about whether and how to engage with messengers offering new information about policy performance elsewhere. They have numerous factors already demanding time and energy as they shape their policies, including past practice and experience, local shareholder desires, supervisor demands, and personal ideology and preferences. In the face of these pressures, a messenger seeking policy-maker time with a promise of more information about others' policy practices seems likely to be ignored.

In such cases, messenger characteristics are of vital importance. In fact, whether individuals accept a particular message may have more to do with external cues than it does with the content of the message itself (Lee 2005, 1001). Scholars have identified credibility as perhaps the most important external cue that makes for communicating persuasive messages (Miller and Krosnick 2000; Druckman 2001; Iyengar and Valentino 2000; Greer 2003; Page, Shapiro, and Dempsey 1987). Despite a scholarly consensus on the importance of credibility, the actual elements that contribute to a source's credibility are contested. Many researchers cite the bias associated with a source as responsible for its credibility. Some have found that the political party and/or ideology

associated with a source affects individual perceptions of the information provided by that source (Baum and Groeling 2008; Druckman 2001; Malka, Krosnick, and Langer 2009; Greer 2003; Pornpitakpan 2004).

In the context of information about development-policy impacts, other source attributes are likely to be more important than ideology. Chief among them is nationality. Some countries have achieved higher levels of development than others, potentially giving them more credibility on development policy. Moreover, some countries have stronger educational systems than others, including at the university level where many impact evaluations are produced. Policy makers seeking high-quality information on the policy practices and outcomes of others are more likely to believe information from countries with strong reputations for institutions of higher education and perceptions of policy effectiveness.

More broadly, studies on the consumption of goods in an international context demonstrate that the reputation of the home country influences on consumer choice. In a review of the literature, Papadopoulos and Heslop (2002, 298) found that country of origin provides a "powerful stereotype" that influences many different sorts of consumers. While it is relatively easy to generate positive feelings toward a given country and a product from that country, it is more difficult to ask consumers to pay a higher price as a result. Yet some scholars have found that consumers are willing to do so (Koschate-Fischer 2012). Fairly often, the country-of-origin effect outweighs even the price or the advertised quality of the product (Okechuku 1994). In a context where policy makers know very little about a particular product, country-of-origin effects may be quite strong. Such effects are especially strong for technological goods (Costa, Carneiro and Goldszmidt 2016). While the service offered in our experiment is not a consumption good as in the above examples, it is plausible that the abstract and technological nature of the policy-evaluation website might demonstrate similar diffusion patterns.

Country brand indices differ on where the United States is placed with respect to other countries, but it is always placed much higher than Peru or other developing countries. According to the Anholt-GFK scale, the United States ranks as number 1 out of 50 countries listed, including various countries from Europe, Asia, Africa, and Latin America (Anholt 2016 http://nation-brands.gfk.com/). Another prominent country branding index ranks the United States as number 7 out of 75 on its index (FutureBrand 2015 http://www.futurebrand.com/country-brand-index). Because in our experiment, all the American researchers not only mentioned their name but also mentioned their country of origin, the United States, the respected "country brand" may have impacted Peruvian government officials' decisions to meet with Americans rather than Peruvians.

Hypothesis 1: Policy makers in developing countries are more likely to seek information from U.S. messengers than from local messengers.

Gender should impact credibility. In many countries, socially constructed and biased gender norms give men more social status vis-a-vis women in issue areas like government decision-making and scientific study. Leadership and power are typically seen as masculine, and there is greater prejudice against women who try to take on positions requiring these characteristics (Garcia-Retamero and López-Zafra 2006). Though the percentage of women in leadership positions has increased significantly since the establishment of descriptive-representation quotas in over one-hundred countries (Beaman et al. 2009), many biases towards female politicians unfortunately persist. UN Women (2018) reports that only 22.8 percent of national parliamentarians were women as of June 2016 and that only 18.3 percent of government ministers were women as of January 2017. In Peru, only 29.7 percent of legislators, senior officials and

managers are female and women hold only 22.3 percent of seats in the national Congress.[2] Even though women may be able to vote and run for office, it is often difficult for them to hold onto positions once elected or to have much power in decision-making (Hazarika and Guha-Khasnobis 2008). More generally, women tend to be undervalued in society, contributing to this preference for male authority.

We hypothesize that policy makers will respond more positively to male messengers, as women may be seen as deviating from traditional gender roles and may be viewed as less credible. Whether consciously or subconsciously, officials may view male researchers as more legitimate and be more likely to express interest in meeting with them. Especially because we are presenting a database of information, which may be risky for government officials to log onto in their offices, they may be even more hesitant to meet with women. Many of these officials are quite busy, and sacrificing even a half an hour of time to a young researcher would require them to believe that the information to be presented will be of benefit to them.

Hypothesis 2: Policy makers in developing countries are more likely to seek information from men than from women.

While messenger characteristics are important, message content also matters. In particular, academics have repeatedly demonstrated that social proofs have important impacts on those receiving a message (Cialdini et al. 1990, Schultz et. al 2008, Shearman 2007). Social proofs refer to messages stating that peers act in particular ways, which influences message recipients to behave in those same ways. Scholars have demonstrated that social proofs operate in a huge variety of

[2] United Nations Development Programme, Human Development Reports, http://hdr.undp.org/en/countries/profiles/PER, accessed 13 Aug. 2018.

contexts, including towel usage in hotels (Schultz 2008), charitable donations (Shearman 2007), and residential water use (Ferraro and Price 2013). Where subjects can observe the behavior of others rather than just hear about it, the effects of social proof or social contagion are also significant, in areas as diverse as sleep, happiness, divorce, sexuality, depression, smoking and tastes in books or movies (Christakis and Fowler 2013). These positive results are not just confined to the United States or Western Europe. In fact, some evidence exists that social proofs may work even better in countries with less individualist-oriented values (Cialdini 1999).

Although few studies focus on political contexts, Gerber, Green and Larimer (2008) found that individuals were far more likely to vote if scholars threatened to publicize whether they went to the polls or not, suggesting the importance of social pressure. In a Facebook experiment with 61 million people that more closely resembles traditional social proof research in psychology, receiving messages that friends had voted significantly increased voter turnout (Bond et al. 2012). Results can even occur in areas where one might not expect much behavioral change. For example, telling those with delinquent taxes that others have paid their taxes resulted in increased payment rates for overdue taxes (Hallsworth et al. 2017).

Hypothesis 3: Policy makers in developing countries are more likely to respond to social proofs than to generic claims about importance and utility.

Research Design

We sent 6- to 8-person teams of student researchers from Brigham Young University to Peru in July and August 2017. Some were masters' students and others were undergraduates, but all were familiar with basic RCT methods and had an interest in development issues. We hired the equivalent number of local researchers with similar qualifications and interests. We provided

each country team with a list of all possible government officials and contact information we could identify through extensive internet searches. We randomly assigned each researcher to an equal number of officials with a blocking algorithm assigning an equal number of officials to each researcher in each government ministry.

Each researcher then contacted the assigned officials by email, phone call, or both.[3] In a few cases, researchers even showed up in person at the official's office unannounced. In contacting the officials, the researchers used their full names. They each had names that were readily identifiable as English-speaking or as local, and as male or female. During the first contact, researchers from the United States and from Peru stated explicitly their nationality to make it even more obvious and to clarify the provenance of the name.

In the social proof condition, researchers told the government official that colleagues in their ministry had already indicated an interest in scheduling a meeting, if that were true. Where it was not true – early in the process before many had indicated an interest – researchers stated that other government officials (without naming a ministry) had indicated an interest. These claims were always true because we began by contacting the control group. Treatment subjects were contacted one or two days later, and all follow-up emails followed a rolling design in which potential temporal effects could be identified (but no significant such effects were identified). In the control condition, researchers claimed simply that academic studies could be important and helpful. The purpose of the initial contact was to invite the official to set up a meeting where the researcher could meet in person to demonstrate our website. The text of the social proof and control paragraphs from our emails follows:

SOCIAL PROOF TREATMENT: A number of your colleagues in [YOUR MINISTRY] have already

[3] The vast majority of contacts and subsequent communications were made by email, and controlling for the fixed effects of phone calls or office visits does not substantively alter the reported results.

expressed interest and have invited a meeting with us. In a recent survey of 6,750 policymakers in 126 countries, information about successful public policy programs in other countries was ranked second in importance on a list of 14 policy inputs.

CONTROL: There are many evidence-based academic studies of policy programs that are important. We think these could be very helpful in guiding your decisions. These studies have taken place all over the world and have covered every important topic area in economic and social development.

Results and Discussion

In total, we attempted to email 2,676 government officials and attempted to call 483 government officials, although some of these emails and phone numbers were invalid. In the following analysis, we include only government officials for whom we did *not* receive a notification that the email was sent to a non-existent address and/or for whom we did not receive a message on the phone telling us that the number was incorrect or that we had the wrong number. In the end, 2,072 emails were sent to valid addresses and 306 phone calls were made to valid numbers, adding up to a total of 2,251 observations in our final analysis (some of the phone calls made were to officials who had already received emails).

We focus on four outcomes from government officials: 1) a positive response to our request for a meeting, 2) a negative response to that request, 3) whether an appointment was held, and 4) how many people were present for the training when an appointment was held. Positive responses refer to any substantive response that indicated an interest in the project while negative responses refer to substantive responses that clearly stated disinterest. Table 1 reports the results. Ministries are not shown for economy of presentation, but ministry fixed effects are included in the regressions as controls.

Table 1. Results of Treatments on Responses of Government Officials

VARIABLES	Positive Response	Negative Response	Meeting Held	Number of Government Officials Trained
Social Proof	-0.0291	0.138*	0.0379	0.0233
	(0.0804)	(0.0824)	(0.100)	(0.0517)
RA from USA	0.435***	-0.652***	0.254**	0.135***
	(0.0621)	(0.0739)	(0.0995)	(0.0381)
Female RA	0.0586	-0.690***	0.0901	0.00759
	(0.0532)	(0.107)	(0.0588)	(0.0241)
Constant	-1.533***	-1.046***	-1.774***	0.0828**
	(0.0783)	(0.0807)	(0.112)	(0.0315)
Observations	2,243	2,243	2,243	2,243
R-squared				0.004

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Nationality of Research Assistant

Positive Response

We found that Peruvian government officials were significantly more likely to respond positively to unsolicited emails and phone calls from U.S. nationals than from Peruvians. Officials gave Peruvians a positive response to 6.43% (95% confidence interval of 4.82 to 8.03) of contacts while researchers from the United States received a positive response 13.90% (95% confidence interval of 11.14 to 16.35) of the time. This difference is statistically significant at the .01 level. Figure 1 (located in the appendix) illustrates how government officials were more than twice as likely to give a positive response to Americans as they were to Peruvians.

Negative Response

Just as government officials were more likely to give a positive response to researchers from the United States, they were also more likely to give negative responses to

Peruvians. Holding all else equal, officials gave negative responses to researchers from the U.S. to 3.29% of the contacts (95% confidence interval of 1.94 to 4.64) while they gave negative responses to Peruvians 11.14% of the time (95% confidence interval of 8.22 to 14.06). This difference is statistically significant at the .01 level. Figure 2 in the appendix illustrates the nearly four-fold increased likelihood that Peruvian researchers had of receiving a negative response over American researchers. Interestingly, this finding runs counter to a vast sea of literature on co-ethnicity preferences in the medical field. For decades, authors in the medical field have shown that patients consult, trust, and prefer doctors of their own race rather than of another race when given a choice (Laveist and Nuru-Jeter 2002). Our research provides evidence that, in the context of receiving policy advice, government officials are less likely to accept an offer for advice and information from locals than from Americans.

Appointment Held

As with other outcomes, Peruvian government officials showed a preference for U.S. messengers over Peruvians in terms of appointments held. All else equal, Peruvians held appointments with 4.33% of officials contacted (95% confidence interval of 2.64 to 6.02), yet Americans held briefings with 7.21% of the officials the contacted (95% confidence interval of 5.77 to 8.66). This difference is statistically significant at the .05 level. The preference for meeting with researchers from the United States is shown in Figure 3.

Number of Government Officials Trained

Researcher nationality also significantly impacted the number of people trained in a meeting. Researchers from the United States trained, on average, 0.133 more people per meeting than their Peruvian counterparts. This result is statistically significant at the .01 level.

Gender of Research Assistant

Positive Response

Against expectations, we did not find any significant difference between the likelihood of male and female researchers receiving a positive response inviting a meeting from a government official. Male researchers received positive responses for 9.79% (95% confidence interval of 7.94 to 11.64) of contacts. Females received positive responses for 10.82% (95% confidence interval of 8.66 to 12.98) of the contacts they made. We illustrate the findings in Figure 4. It must be noted here that the difference-in-means results contrast with the regression coefficients in a critical way. Due to vagaries in researcher recruitment, a greater proportion of American researchers were women than Peruvians, creating a randomization imbalance that needs adjustment using regression. While the difference in means suggests a significant difference in favor of female researchers, the adjusted coefficients in regression analysis fail to suggest a meaningful gender difference in messenger appeal to officials. These results suggest that the expected gender bias, despite being obviously present in other aspects of Peruvian politics and government, did not manifest in this context in a way that can be detected statistically.

Negative Response

Female researchers were found to be significantly less likely to receive a negative response from a government official to the invitation to hold a meeting. Holding all else equal, male researchers received negative responses to 10.64% (95% confidence interval of 8.16 to 13.12) of contacts. Female researchers only received negative responses to 2.85% (95% confidence interval of 1.14 to 4.57) of their inquiries. This difference is statistically significant at the .01 level. Results are illustrated in Figure 5. Again, this finding is against expectations but is also significant

statistically.[4]

Appointment Held

Although female researchers were somewhat more likely to hold a meeting with government officials than male researchers, this result was not statistically significant. Male researchers held appointments with 5.34% (95% confidence interval of 4.13 to 6.54) of contacted officials. Female researchers, on the other hand, provided actual briefings to 6.38% (95% confidence interval of 4.81 to 7.95) of their contacts. These results are illustrated in Figure 6.

Number of Government Officials Trained

The gender of the research assistant had no significant effect on the number of people in attendance at the meeting.

Discussion

We probed these results to determine whether they were driven by the gender of the government official. We had no expectations of whether female or male government officials might have preferred women researchers. In marketing research, customers generally find product endorsers of the same gender as more trustworthy (Boyd and Shank 2004). At the same time, it is possible that anticipated female attractiveness played some role in male preferences to interact with women. In an experiment soliciting charitable donations, Landry et al. (2006) found that donations increased by 35 to 72 percent with a one standard-deviation increase in physical attractiveness level of a female solicitor. No significant increase in donations was observed when

[4] We are still validating this data. We think one female RA may have significantly mis-reported the number of negative responses received.

varying the physical attractiveness level of a male solicitor. This result was driven almost entirely by households where men answered the door. In our experiments, government officials knew nothing about researcher attractiveness because our first contacts were made by email or telephone. Still, they may have anticipated relatively attractive researchers coming to their offices if they were female names. After analyzing the data for these possibilities, however, we found no results that are inconsistent with the null hypothesis of zero interaction effect.

What then explains the results for the one outcome – analyzing negative responses – favoring female researchers? One possibility is that women are seen as more trustworthy across various social contexts. In the broader literature, findings on gender and trust are somewhat mixed, though women are sometimes perceived to be more trustworthy and in fact often behave in more trustworthy ways (Slonim and Guillen 2010). In a widely used investment game, senders decide how much of a given amount of money to send to a recipient. The money is then tripled in value by the experimenters, and the recipient decides how much to send back. Buchan, Croson and Solnick (2008) found that women rewarded the trust placed in them as recipients by giving more back. In other studies focusing on unsolicited door-to-door fundraising, much like the experiment that we ran consisting of cold contact asking people for their time, women received significantly more donations/success than men (Landry et al. 2006).

In a sophisticated version of the game, Slonim and Guillen (2010) found that both men and women prefer to select recipients of the opposite gender as their partners if given the chance. However, both men and women believed that women were more trustworthy than men. Men sent significantly more money to women, but women did not send significantly more money to men. The authors found that these patterns are explained by the tendency for men to prefer to interact with women and for men to trust women more. Other experimental research on donation generosity shows that female recipients receive more money than male recipients (Dufwenberg and Muren

2006). Perhaps Peruvian government officials view women as more trustworthy and thus responded less negatively to them and more negatively to men. Still, this one result is not replicated for the other dependent variables and, compared to positive responses, proves less relevant to the downstream outcomes of appointments held and number of people trained.

The Social Proof Treatment

Research has consistently shown social proofs to be an effective way to influence the decisions of people (Schultz et. al 2008, Shearman 2007). In our experiment, however, we found that social proof had, at best, no significant effect on response and website usage rates and may have even had a negative effect on some outcomes, although this effect is not significant statistically for any of the four dependent variables.

Positive Response

The social proof treatment resulted in a slightly lower average likelihood of a government official responding positively, although this decrease in probability is not statistically significant. Government officials assigned to the control group responded, on average, to 10.54% (95% confidence interval of 7.99 to 13.10) of invitations to hold an appointment to learn more about using evidence-based reports in their policy work. Those assigned to the social proof treatment responded positively at the rate of 10.03% (95% confidence interval of 8.15 to 11.91). We illustrate the findings in Figure 7. These results are surprising, as social proofs have generally been found to significantly increase uptake rates for behaviors of peers (Cialdini et al. 1990).

Negative Response

The social proof treatment also slightly increased the probability of a government official

responding negatively. Government officials that did not receive the social proof treatment had a predicted 6.81% probability (95% confidence interval of 4.83 to 8.78) of receiving a negative response for every government official contacted. Those who were randomly assigned to the social proof treatment, however, had an 8.59% probability (95% confidence interval of 6.06 to 11.12) of receiving a negative response. This difference is significant at the .1 level. Results are illustrated in Figure 8.

Appointment Held

Contrary to the above results, government officials who were assigned to the social proof treatment were slightly more likely to hold an appointment with a RA (though, again, this difference in probability was not statistically significant). Government officials in the control group on average held meetings in 5.63% (95% confidence interval of 4.06 to 7.20) of researcher contacts while government officials assigned to the social proof treatment had contacts result in briefings 6.07% of the time (95% confidence interval of 4.32 to 7.81). Figure 9 presents graphically these results.

Number of Government Officials Trained

Contrary to all previous finding about the social proof, the treatment slightly increased the predicted number of government officials trained in a meeting, though this result was not significant statistically.

Discussion

While abundant literature is available regarding social proofs, little research indicates that social proofs are ineffective. What might explain our findings? The results are especially

puzzling when considering that our treatments were quite similar to other social-proof treatments in the literature. In particular, the statement was descriptive – stating what other people do – rather than injunctive – stating what society generally approves or disapproves. Descriptive norms have been shown to be more effective at influencing behavior (Jacobson, Mortensen, and Cialdini 2011). We were careful not to normalize any undesirable behavior, such as implying that many government officials do not use academic studies, as social proofs that normalize undesirable behavior are often less effective in inspiring the desired action from individuals (Cialdini et. al 2006).

Lastly, we attempted to make the social proof as impactful as possible by stating, whenever true, that a policymakers' colleagues from their own ministry had invited a meeting with researchers. Studies have shown that knowing that similar people are involved in an activity increases one's likeliness of participating (Bond 2012). In cases where we had no scheduled meetings with employees of a certain ministry, we simply stated that other government officials from their country had invited a meeting with us.

One possible explanation of the null result is that elites react differently to social proofs than others might. Few social proof studies focus on elites and even fewer on elites in government positions (Rao 2001). Previous studies indicate that persons classified as elite react differently when receiving advice and making decisions, and this could potentially explain their apparent resistance to the effect of a social proof (Galinsky et. al 2008). Elites have been shown to be less likely to listen to the advice of others and more likely to make decisions alone and on impulse (Fast et. al 2012, See et. al 2011). This could explain why the subjects of our study did not react positively to the social proof treatment. Rather than being motivated by the decisions of others, elites choose to form their own opinions and make their own decisions.

Another possibility is that social proof in the workplace simply induces free-riding. If

employees hear that others in their ministry have received the same training, they may think it less important to spend their time doing it. They could rationally believe that their coworkers or supervisors will bring the issue to their attention later if it is actually important.

Conclusion

What accounts for policy makers' interest in new information about policy efforts and outcomes in other countries? What factors influence the policy diffusion process at a micro-foundational level in developing countries? Our results suggest that the nationality and gender of the messenger have important influences. In Peru, U.S. messengers had much more success than local messengers. We expect that messenger credibility was essential. One important mechanism in the diffusion of policy ideas may well relate to the national origin of the idea or technology. Ideas from some countries may be more likely to diffuse than from others. Contrary to our expectations, women messengers had fewer negative responses, though this result did not replicate across the other, more important, outcomes. Our findings suggest that government officials respond not at all to the social pressure of their peers, and that such subtle influence may even backfire. This finding stands in sharp contrast to volumes of studies about social proofs, which show that they work in many different circumstances. We can only speculate as to why the social pressure intervention failed.

On the whole, the experiment provides a proof of concept for testing key mechanisms in how ideas diffuse from one country to another. We tested three possible causes of openness to diffusion: credibility as manifest in messenger nationality, credibility by way of messenger gender, and social proof in terms of signals of peers' behavior. Only messenger nationality demonstrated evidence of effectiveness in heightening openness to diffusion inception. These results suggest that the content of ideas cannot be easily separated from their packaging, and a

key element of the package is the nationality of messenger herself. More research will need to be done to probe the robustness of the finding as well as to explore other aspects of credibility and persuasiveness. Additionally, we will need to do more to understand if the strong U.S.-nationality result is confined to a bilateral U.S.-Peru dyad or extends more generally to perceptions of U.S. credibility in Latin America broadly or to the developing world at large. Moreover, is it a U.S.-specific effect, or would messengers also prove effective if they hailed from other advanced industrial economies that are home to successful research endeavors.

Finally, it is worth underscoring that this represents more than an academic exercise. The study involves a sincere attempt to curate the vast RCT literature in development policy and programs, simplify it conceptually, and make it wholly accessible through short summaries and consistent, comparable graphical representations. The search for interventions that heighten uptake by the actual officials who are the targets of interest is therefore both relevant to academic discussions about diffusion mechanisms but also a means of learning how to spread rigorous policy knowledge most effectively.

References

Anholt, Simon. 2016. Anholt-GFK Nation Brands Index. http://nation-brands.gfk.com/ accessed September 28, 2017.

Baum, Matthew A. and Tim Groeling. 2008. "New Media and the Polarization of American Political Discourse." *Political Communication* 25, no. 4 (November): 345-365.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly Journal of Economics* 124, no. 4 (November): 1497-1540.

Bond, Robert M., Christopher J. Farris, Jason J. Jones and James A. Fowler. 2012. *Nature* 489 no. 7215: 295-298.

Boyd, TC, and Shank MD. 2004. Athletes and Product Endorsers: The Effect of Gender and Product Relatedness. *Sports Marketing Quarterly* 13, no. 2.

Buchan, Nancy R., Rachel T.A. Croson, and Sara Solnick. 2008. "Trust and Gender: An Examination of Behavior and Beliefs in the Investment Game." *Journal of Economic Behavior and Organization* 68 (3-4): 466-476.

Butler, Daniel M., Craig Volden, Adam M. Dynes, and Boris Shor. 2017. "Ideology, Learning, and Policy Diffusion." *American Journal of Political Science* 61, no. 1 (January): 37-49.

Cairney, Paul, Kathryn Oliver, and Adam Wellstead. 2016. "To Bridge the Divide between Evidence and Policy: Reduce Ambiguity as Much as Uncertainty." *Public Administration Review* 76, no. 3: 339-402.

Christakis, Nicholas A. and James H. Fowler. 2013. "Social Contagion Theory: Examining Dynamic Social Networks and Human Behavior." *Statistics in Medicine* 32: 556-577.

Cialdini, Robert B., Linda J. Demaine , Brad J. Sagarin , Daniel W. Barrett , Kelton Rhoads and Patricia L. Winter. 2006. Managing Social Norms for Persuasive Impact. *Social Influence* 1 no. 1: 3-15.

Costa, Camila, Jorge Carneiro, Rafael Goldszmidt. 2016. "Country Image Effect on Product Assessment: Moderating Role of Consumer Nationality." *Review of Business Management* 18, no. 59 (January/March): 24-42.

Deaton, Angus. 2013. *The Great Escape: Health, Wealth, and the Origins of Inequality*. Princeton, NJ: Princeton University Press.

Druckman, James N. 2001. "The Implications of Framing Effects for Citizen Competence." *Political Behavior* 23, no. 3 (September): 225-256.

Dufwenberg, M., & Muren, A. 2006. Generosity, anonymity, gender. *Journal of Economic Behavior & Organization* 61, no.1: 42-49.

Fast, Nathanael J., Niro Sivanathan, Nicole D. Mayer and Adam D.Galinsky. Power and Overconfidence Decision-Making. *Organizational Behavior and Human Decision Processes* 117 no. 2 (March): 249-260.

Ferraro, Paul J. and Michael K. Price. 2013. "Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiement." *The Review of Economics and Statistics* 95, no. 1 (March): 64-73.

FutureBrand. "Country Brand Index." FutureBrand. https://www.futurebrand.com/country-brand-index (accessed August 13, 2018).

Galinsky, Adam D., Joe C. Magee, Deborah H. Gruenfeld, Jennifer A. Whitson, Katie A. Liljenquist. Power Reduces the Press of the Situation: Implications for Creativity, Conformity, and Dissonance. 2008. *Journal of Personality and Social Psychology* 95, no. 6: 1450-1466.

Garcia-Retamero, Rocio, Esther Lopez-Zafra. 2006. "Prejudice Against Women in Male-Congenial Environments: Perceptions of Gender Role Congruity in Leadership." *Sex Roles* 55, no. 1-2 (July): 51-61.

Gerber, Alan S., Donald P. Green, and Chistopher W. Larimer. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review* 102, no. 1 (February): 33-48.

Graham, Erin R., Charles R. Shipan, and Craig Volden. 2013. "Review Article: The Diffusion of Policy Diffusion Research in Political Science." *British Journal of Political Science* 43, no. 3 (July): 673-701.

Greer, Jennifer D. 2003. "Evaluating the Credibility of Online Information: A Test of Source and Advertising Influence." *Mass Communication and Society* 6, no. 1: 11-28.

Hallsworth, Michael, John A. List, Robert D. Metcalfe, and Ivo Vlaev. 2017. "The Behavioralist as a Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance." *Journal of Public Economics* 148 (April): 14-31.

Head, Brian W. 2015. "Relationships between Policy Academics and Public Servants: Learning at a Distance?" *Australian Journal of Public Administration* 74, no. 1: 5-12.

Head, Brian W. 2016. "Toward More 'Evidence-Informed' Policy Making?" *Public Administration Review* 76, no. 3: 472-484.

Iyengar, Shanto and Nicholas A. Valentino. "Who Says What? Source Credibility as a Mediator of Campaign Advertising." In *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*, edited by Lupia, Arthur, Matthew D. McCubbins, and

Samuel L. Popkin. Cambridge: Cambridge University Press, 2000.

Jacobson, RP, Chad R. Mortensen and Robert B. Cialdini. 2011. Bodies Obliged and Unbound: Diferentiated Response Tendencies for Injunctive and Descriptive Social Norms. *Journal of Personality and Social Psychology* 100 no. 3: 433-448.

Kenny, Charles. 2011. *Getting Better: Why Global Development is Succeeding – And How We Can Improve the World Even More*. New York: Basic Books.

Koschate-Fischer, Nicole, Adamantios Diamantopoulos, and Katharina Oldenkotte. 2012. *Journal of International Marketing* 20, no. 1: 19-41.

Landry, C. E., Lange, A., List, J. A., Price, M. K., & Rupp, N. G. 2006. Toward an understanding of the economics of charity: Evidence from a field experiment. *The Quarterly journal of economics* 121, no 2: 747-782.

Laveist, Thomas A and Amani Nuru-Jeter. 2002. Is Doctor-Patients Race Concordance Associated with Greater Satisfaction with Care? *Journal of Health and Social Behavior* 43, no. 3 (September): 296-306.

Maggetti, Martino and Fabrizio Gilardi. 2016. "Problems (and Solutions) in the Measurement of Policy Diffusion Mechanisms." *Journal of Public Policy* 36, no. 1 (March): 87-107.

Malka, Ariel, John A. Krosnick, and Gary Langer. 2009. "The Association of Knowledge with Concern About Global Warming: Trusted Information Sources Shape Public Thinking." *Risk Analysis* 29, no. 5 (May): 633-647.

Miller, Joanne M. and Jon A. Krosnick. 2000. "News Media Impact on the Ingredients of Presidential Evaluations: Politically Knowledgeable Citizens are Guided by a Trusted Source." *American Journal of Political Science* 44, no. 2 (January): 301-315.

Newman, Joshua, Adrian Cherney, and Brian W. Head. 2017. "Policy Capacity and Evidence-Based Policy in the Public Service." *Public Management Review* 19, no. 2: 157-174.

Okechuku, Chike. 1994. The Importance of Product Country of Origin: A Conjoint Analysis of the United States, Canada, Germany and the Netherlands. *European Journal of Marketing* 28 no. 4: 5-19.

Papadopoulos, N. and L. Helsop. 2002. "Country Equity and Country Branding: Problems and Prospects." *Journal of Brand Management* 9, no. 4 (April): 294-314.

Page, Benjamin I., Robert Y. Shapiro, and Glenn R. Dempsey. 1987. "What Moves Public Opinion?" *The American Political Science Review* 81, no. 1 (March): 23-44.

Pinker, Steven. 2018. *Enlightenment Now: The Case for Reason, Science, Humanism, and Progress*. New York: Penguin.

Pornpitakpan, Chanthika. 2004. "The Persuasiveness of Source Credibility: A Critical Review of Five Decades' Evidence." *Journal of Applied Social Psychology* 32, no. 2: 243-281.

Rao, Hayagreeva, Henrich R. Greve and Gerald F. Davis. 2001. Fool's Gold: Social Proof in the Initiation and Abandonment of Coverage by Wall Street Analysts. *Administrative Science Quarterly* 46 no. 3 (September): 502-526.

Rosling, Hans, Anna Rosling Rönnlund, and Ola Rosling. 2018. *Factfulness: Ten Reasons We're Wrong about the World – And Why Things are Better Than You Think*. New York: Flatiron Books.

Schultz, Wesley P., Azar M. Khazian and Adam C. Zaleski. 2008. Using Normative Social Influence to Promote Conservation Among Hotel Guests. *Taylor and Francis Online* 3, no. 1: 4-23.

See, Kelly E., Elizabeth W. Morrison, Naomi B. Rothman and Jack B. Soll. 2011. The Detrimental Effects of Power on Overconfidence, Advice Taking, and Accracy. *Organizational Behavior and Human Decision Processes* 116, no. 2 (November): 272-285.

Shearman, Sachiyo M. 2007. "Even a Penny Will Help!": Legitimization of Paltry Donation and Social Proof in Soliciting Donation to a Charitable Organization. *Taylor and Francis Online* 32, no. 4: 556-577.

Simmons, Beth A., Frank Dobbin and Geoffrey Garrett. 2006. "Introduction: The International Diffusion of Liberalism." *International Organization* 60, no. 4 (October): 781-810.

Slonim, Robert and Pablo Guillen. "Gender Selection Discrimination: Evidence from a Trust Game." *Jornal of Economic Behavior & Organization* 76, no. 2: 385-405.

UN Development Programme. "Human Development Reports." United Nations. http://hdr.undp.org/en/2016-report (accessed August 13, 2018).

UN Women. "Women's Leadership and Political Participation." United Nations. http://www.unwomen.org/en/what-we-do/leadership-and-political-participation/facts-and-figures (accessed August 13, 2018).

Figure 1: Nationality and Probability of Positive Response

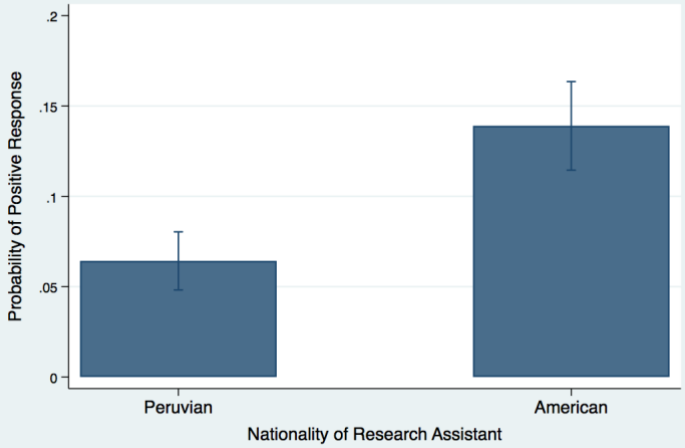

Figure 2: Nationality and Probability of Negative Response

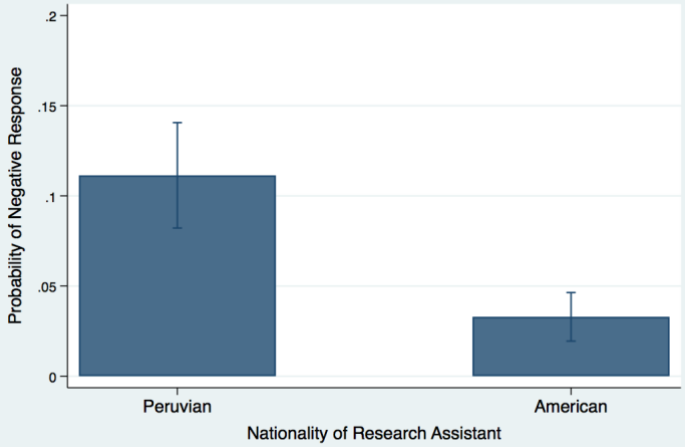

Figure 3: Nationality and Probability of Holding Appointment

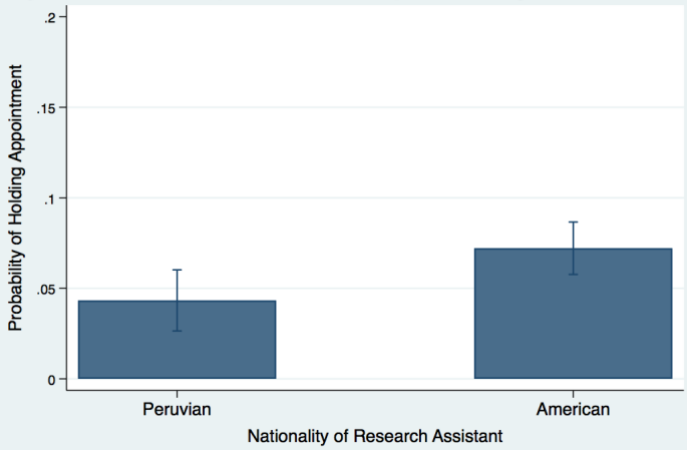

Figure 4: Gender and Probability of Positive Response

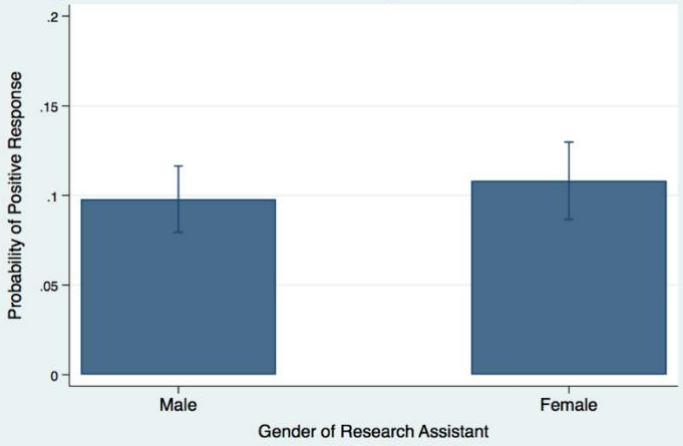

Figure 5: Gender and Probability of Negative Response

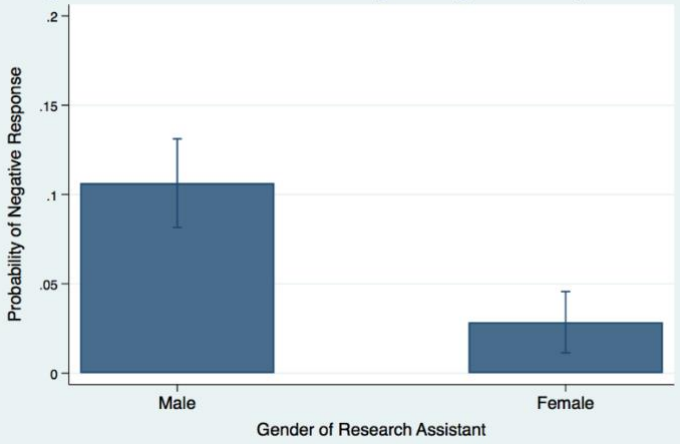

Figure 6: Gender and Probability of Holding Appointment

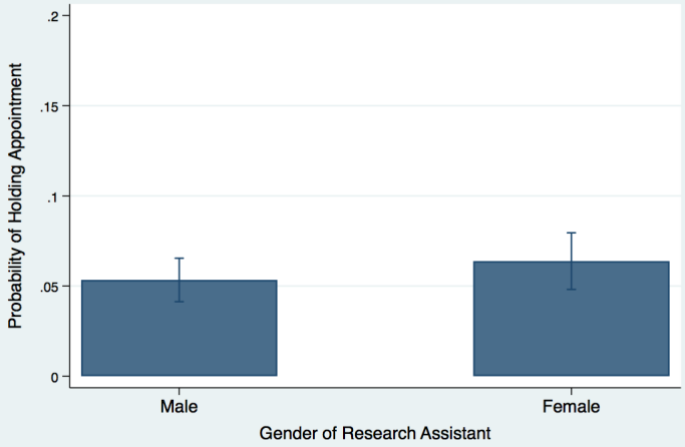

Figure 7: Social Proof and Probability of Positive Response

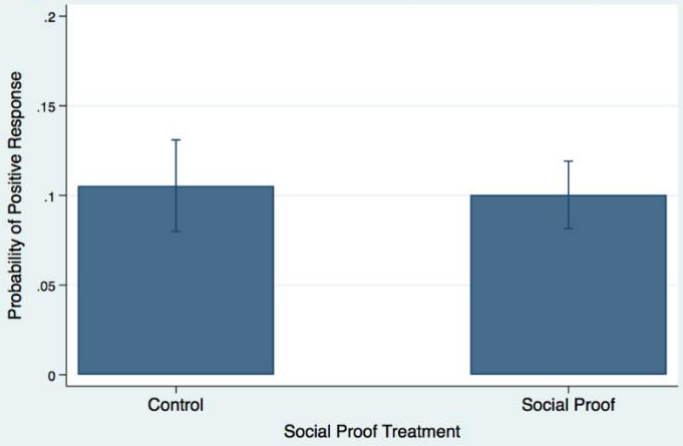

Figure 8: Social Proof and Probability of Negative Response

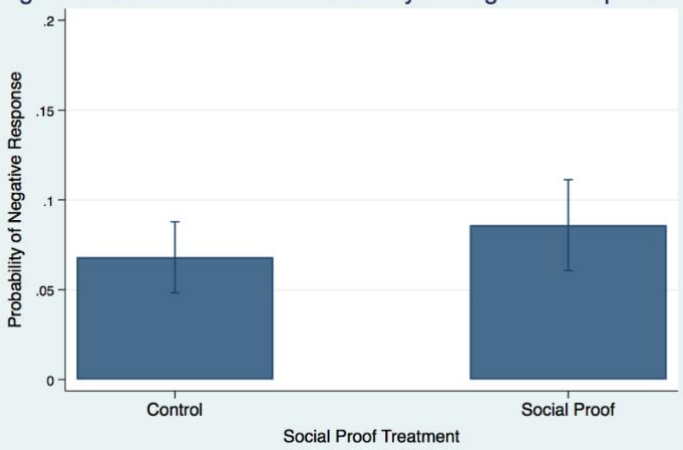

Figure 9: Social Proof and Probability of Holding Appointment