Received: 13 July 2022 | Revised: 28 April 2023 | Accepted: 2 May 2023

DOI: 10.1111/rode.13008

REGULAR ARTICLE

WILEY

International or national implementers—Who is better? Evidence from a framing experiment

Lennart C. Kaplan[1,2,3] | **Jana Kuhnt**[1] | **Katharina Richert**[4,5]

[1]German Institute of Development and Sustainability, Bonn, Germany

[2]University of Göttingen, Göttingen, Germany

[3]Kiel Institute for the World Economy, Kiel, Germany

[4]Center for Evaluation and Development, Mannheim, Germany

[5]University of Mannheim, Mannheim, Germany

Correspondence
Jana Kuhnt, German Institute of Development and Sustainability, Tulpenfeld 6, 53113 Bonn, Germany.
Email: jana.kuhnt@idos-research.de

Funding information
European Commission, Grant/Award Number: Experts4Asia; German Research Foundation, Grant/Award Number: RTG-1723; Lower Saxonian Ministry of Science and Culture, Grant/Award Number: "Reducing Poverty Risk" project

Abstract

Behavioral economics has shown that changing small features in framing a context or action may drastically change behavior. A key factor characterizing most development interventions is the salience of either a local or an international implementer. Using the setup of an intervention conducted in Indonesia, we show that the study population in the Acehnese context exhibits higher levels of support for the project if the participation of international actors is highlighted. We find that previous experience with the respective actor is pivotal. Qualitative evidence suggests that internationals' perceived skills drive results, highlighting the importance of strengthened local capacities for positive experiences with local implementers. Overall, the study underlines the benefits of linking framing experiments to the actual experiences of respondents to generate insights into the real world.

KEYWORDS

behavioral economics, framing experiment, public health

1 | INTRODUCTION

A large focus in the literature studying development cooperation naturally lies on its effectiveness (Burnside & Dollar, 2000; Easterly et al., 2004). Support and uptake by the target population are major factors influencing the success of an intervention. Behavioral economics stresses in this regard the importance of non-monetary incentives that shape human motivation and

behavior and consequently, the successful design of interventions (e.g., Ashraf et al., 2014; Banerjee et al., 2010; Bowles & Polania-Reyes, 2012; Cole et al., 2013; Gneezy et al., 2011). One of the core features of a project, which may affect the support of development programs, is typically the identity of the implementing team. Recent research stresses the importance of the implementing agent for the behavior of the target group. Higher support for a specific group of implementers could be driven by heuristics or behavioral biases, for example, stereotypes. However, these are usually based on underlying perceptions regarding the implementers. One might not a priori prefer international or local implementers, but support those known for higher implementation capacities, for instance. Dietrich and Winters (2015) and Winters et al. (2017) show that respondents link higher quality perceptions to donors rather than the national government. Additionally, Findley et al. (2017) and Milner et al. (2016) demonstrate that respondents perceive international actors to be less entrenched in the local political economy and, hence, perform better with regard to curbing corruption. Another reason might solely be the higher visibility among international donors (Vollmer, 2012), which is ultimately targeted to affect recipients' perceptions. In contrast, the "home bias"-phenomenon suggests that cultural proximity could increase people's trust with local agents (e.g., Fuchs & Gehring, 2017), where particularly for development programs, local implementers may offer tangible benefits of context knowledge. Cilliers et al. (2015) show that the presence of a foreigner versus a local as a third-party bystander positively affects the contributions of participants in a dictator game in Sierra Leone and identify two potential channels: First, an increase in contributions to impress the foreigner and, second, reduced contributions in areas that were previously exposed to development cooperation projects. In the latter locations, they show that participants more frequently believed that the game tested their need for aid, and subsequently contributed less. The previous exposure (here with aid) is shown to be an important factor in shaping perceptions, attitudes, and subsequent support for projects.

We add to this literature and deepen it by linking the implementers' identity to beneficiaries' actual work experience with those actors. For this, we make use of the introduction of the World Health Organization (WHO)'s Safe Childbirth Checklist (SCC) in Indonesia's Aceh province (Kaplan et al., 2021). Two-thirds of maternal and newborn deaths globally occur due to causes, which could largely be prevented if well-established essential practices were followed (WHO, 2018). The WHO developed the SCC, a four-page checklist to address the major risk factors for mothers and children around birth.[1] Despite a high commitment of healthcare providers during the launch events of the SCC, uptake was lower than expected (see also Figures C2 and C3). We realized during our research that local association of the implementing agents is likely to influence behavior of the target group toward the project. By conducting a framing experiment among the health staff in a real-world setting, we causally assess whether health personnel's support toward checklist use changes conditional on whether the participation of local or international agents in the project is highlighted.

Our results indicate that the change in support for the SCC project is due to the salience of international versus local involvement. The population under study shows greater support for interventions with international involvement. Previous exposure to both international and local implementers drives those positive behavioral reactions toward international research projects. A complementary qualitative analysis suggests that a more favorable assessment of international agents' skills drives results. Increasing the visibility and branding of international actors is a low-hanging fruit to encourage take-up. However, for a more sustainable implementation in the long-run, one should investigate how strengthened local capabilities can create positive experiences with local actors and ultimately increase project support.

The study is structured as follows: Section 2 describes the background of our study. Section 3 describes our research design and data. Section 4 elaborates on the methods used, and the results are described in Section 5. Section 6 discusses the generalizability and policy relevance of the results and concludes the study.

2 | STUDY BACKGROUND

After 30 years of civil war and the 2004 Indian Ocean tsunami, Aceh was subject to massive reconstruction efforts by the national government and international donors (Doocy et al., 2007), but the province still ranks comparatively low concerning addressing neonatal and maternal mortality (Diba et al., 2019). The WHO SCC entails the essential practices addressing the major risk factors for mothers and children in low- and middle-income countries. Experience from other medical fields suggests checklists to be a promising tool to motivate health personnel to follow essential practices and tackle the know-do gap. This gap between the knowledge about what should be done to ensure safe deliveries and what is actually done is large. Insights from behavioral economics suggest that human behavior is bounded by limitations of the working memory. In situations characterized by high levels of cognitive load—the amount of mental activity imposed—the successful execution of certain tasks might be interrupted or impaired (e.g., Burgess, 2010; Deck & Jahedi, 2015; Lichand & Mani, 2016). Checklists can be especially helpful to reduce additional cognitive load and allow a reduction of complexity of the situation at hand by reminding the user of the essential steps to follow (e.g., Borchard et al., 2012; Haugen et al., 2015; Workman et al., 2007). In that respect, checklists could be key to much needed efficiency improvements in the health systems of low- and middle-income countries (Grigoli & Kapsoli, 2018).

The original WHO SCC study used a cluster randomized control design to rigorously evaluate the effect of the SCC on maternal and newborn health. The international research team implemented the checklist with a light-touch approach in collaboration with local partners.[2] Due to the post-tsunami reconstruction work, Aceh was exposed to various local and international projects, which facilitates the assessment of the different implementers in the given context.

3 | RESEARCH DESIGN

To study the impact of the salience of international versus local project implementers on the target group's support, we make use of the treatment-control design of the original SCC evaluation. Specifically, we conduct a framing experiment with midwives in the control group to avoid the framing is contaminated by heterogeneous experience with checklist usage. The framing experiment is, thus, embedded in the original study setting as described in Figure 1.

Stressing certain attributes of a particular situation among otherwise equivalent descriptions can lead to very different perceptions and behavioral reactions (Johnson & Goldstein, 2003; Kahneman, 2003; Tversky & Kahneman, 1981). This is evidenced across income settings (Banuri et al., 2019; Hong et al., 2015). The result is what is called the *framing effect* (Tversky & Kahneman, 1980).[3]

Framing experiments are a valuable tool to generate policy-relevant insights in order to understand the underlying structural mechanisms (Duflo et al., 2007; Viceisza, 2015). We

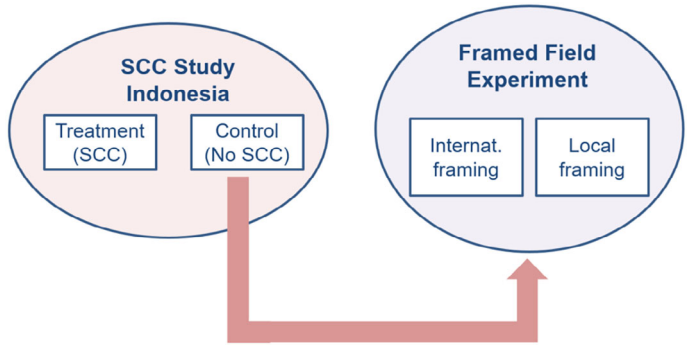

FIGURE 1 Study design flow chart. *Source*: Authors' depiction. [Colour figure can be viewed at wileyonlinelibrary.com]

consider the following question: "Everything equal, how likely are healthcare providers to support the intervention given that the research and implementation team is international or local?" Evidently, an effective framing treatment asks for the respondents not to be aware of the de facto identity of implementers.

Within the group of health practitioners working at Indonesian control facilities (of the original WHO SCC study), we used a between-subject design and randomly assigned the study participants to two different framing information related to the actually conducted SCC intervention.[4]

The first framing information stressed the involvement of international actors in the intervention, while the second made the participation of local counterparts more salient.[5]

We conducted the experiment in total with 236 female midwives. In a short pre-experimental survey, we collected background information, including socio-economic and contextual work characteristics, of each participant.[6] In appreciation for their survey participation, each respondent received a voucher for a phone credit top-up worth 25,000 IDR (approx. 1.75 US$). Afterward, the enumerators offered the respondents to participate in the experiment.[7]

The "experimental commodity" was derived from the ongoing original SCC intervention. First, the idea and structure of the SCC were explained to the participants. Afterward, they were presented with one of the two framings that selectively either stressed the involvement of "local" or "international" actors respectively, in the SCC intervention. We used the fact that the SCC evaluation has been implemented jointly by both—international and local—actors and, therefore, highlighted different attributes of the project. Lastly, we conducted a short post-experimental survey, including questions capturing potential framing mechanisms and additional control variables, like the experience of current financial distress (Box 1).

We then investigated the participants' respective behavior toward the intervention by assessing the support for the SCC project. We proxy SCC support by asking the respondents whether they would contribute to buy checklist copies, which would support the implementation of the SCC in other anonymous health facilities within the province.[8]

The monetary contribution was directly deducted from the voucher for phone credit top-up in appreciation of their survey participation.[9] The contribution was made anonymously. To create transparency on the use of the collected funds, we publicly made information on total amounts available after the end of the study and informed the participant about this procedure. Further, to counter potential bias through speculations on the financial capabilities of different

BOX 1 Framing

"Among other researchers, *[INTERNATIONAL/LOCAL]* researchers took an active role in introducing the checklist to 17 facilities in Aceh province. The research team received approval from the provincial health office of Aceh. However, no funding was provided by the provincial health office. *[LOCAL/INTERNATIONAL]* research assistants and *[INTERNATIONAL/LOCAL]* health professionals with a lot of experience in delivery services were important partners and greatly supported the project. I will now read to you information about the funding of the Safe Childbirth study conducted by the *[INTERNATIONAL/LOCAL]* researchers." For the full experimental protocol please see Appendix A1.

actors, we stress that funding of the intervention is ensured irrespective of the framing information given to the participant.

In the post-experimental survey, we asked several questions on potential mechanisms to explain differential preferences toward implementers. These questions related to perceived corruption, sufficient funding capabilities, accountability, skills, and control to implement interventions. All this data was collected after the experiment to not affect our main outcome measures. However, this procedure comes with the trade-off of potential justification bias, where individuals would adapt their answers ex-post to justify the previously indicated support. We indeed find that the framing statistically significantly affects some of these variables.[10] We did not use those channels for further analysis.

In order to get a clearer understanding how previous experiences with local and international project implementers affect perceptions, we conducted a follow-up open-ended qualitative survey. In those surveys, we asked "In your opinion, what are some of the strengths and challenges of international projects?" and "Please describe your experience working with international teams." Answers complement the findings on experience with local and international agents.

4 | EMPIRICAL APPROACH

Our analysis of the framing experiment aims to identify the existence of a systematic difference in the support for our intervention by health practitioners, conditional on whether the local or international implementation was more salient. Since we randomized participants into different treatment groups, we can make causal inference on how the origin of implementers affects indicated support for the SCC intervention. Our results are based on the following regression equation:

$$y_i = \alpha + \beta_1 \text{framing}_i + \beta_m \sum_m C_i + v_i \quad (1)$$

In our most parsimonious model, y_i is the outcome variable, indicating the support of the SCC by health worker i. α is a constant, and framing_i is a binary variable, which equals one if

the respondent was exposed to an international, and zero for a local framing. We are, thus, mainly interested in the effect size of β_1.

In adjusted regressions, we add $\sum_m C_i$, which is our set of control variables. The controls include a variable indicating the respective facility type, where the participant is employed. Moreover, we add a continuous indicator of perceived resource access at health facilities. We chose to focus on those two variables since there is a statistical difference across framings, which otherwise may induce biased estimates.[11] Above those two variables, we also adjust for the facility-level cesarean section rate as a measure for risk births, given that we learned from qualitative interviews that SCC usage was lower during emergencies. Finally, v_i describes the residual. Errors are clustered at the facility level to take into account similarities within teams. In the final set of regressions, we consider heterogeneous effects of the framing via an interaction term.

$$y_i = \alpha + \beta_1 \text{framing}_i + \beta_2 \text{framing}_i^* c_i + \beta_3 c_i + \beta_m \sum_m C_i + v_i, \quad (2)$$

where c_i is previous participation in international or local projects. While the randomization ensures exogeneity of the framing, project participation is potentially endogenous regarding other traits of the surveyed respondent. However, as methodological research by Bun and Harrison (2018) and Nizalova and Murtazashvili (2016) indicates, the interaction of an exogenous and an endogenous variable can be considered as exogenous, when controlling for the endogenous variable.[12]

5 | RESULTS

5.1 | Descriptives

In our main analysis, we focus on those participants that have not been in prior contact with the SCC as 27.92% of the respondents state that they were previously exposed to the SCC.[13] As we cannot infer how much these respondents know about the SCC intervention and how intense the exposure was, excluding them is the more conservative choice.[14]

This reduces our sample to 170 participants.[15] Individual characteristics and further contextual variables are generally balanced across framings while some facility-level variables (facility type and resource access) are unbalanced (Tables D1 and D2). This is considered in the subsequent analysis. Previous SCC exposure was equally distributed across the framing treatments, ruling out selection concerns and enabling us to interpret the estimates causally.

In the post-experimental survey, we asked participants whether they have previously participated in interventions by international or local experts or researchers, respectively. In the Acehnese health sector, 10% (17.5%) of the surveyed providers have previously participated in research projects by international (local) actors. Those interactions date back significantly before our intervention as only 2.5% of the respondents faced international research projects in their facility during the previous 2 years.

5.2 | Main results

Table 1 displays the main results of the framing experiment conducted in Indonesia, which build on Equation (1).[16] The first column presents the unadjusted results, whereas the second

TABLE 1 Framing experiment—Main results.

Financial contribution in support of SCC project (in IDR)			
	(a)	**(b)**	**(c)**
Framing: 1 = "international"	756	1346*	1394*
p-value	.290	.065	.058
RI *p*-value	.326	.107	.08
Pub. Hospital		−3647** (.017)	−3246** (.045)
Priv. Hospital		−2561 (.302)	−1.584 (.651)
Resource Access		−962 (.295)	−667 (.453)
C-Sec. Rate			−32 (0.458)
N	167	167	167
Control variables	No	Yes	Yes
Mean of dep. var.	4758	4758	4758
SD of dep. var.	4711	4711	4711

Note: All specifications are based upon the sample limited to those respondents without prior Safe Childbirth Checklist (SCC) contact. The same regression with wild cluster bootstrapped SE can be found in Appendix Table D3, for which significance levels hold. RI *p* values are computed with a permutation test based on Hess (2017). Asterisks indicate *p* values based on standard errors clustered at the facility level: *$p < .1$, **$p < .05$, *** $p < .01$.

column gives the results adjusted for additional control variables, which we found to be unbalanced in the experimental balance test. We further add in the third column, the rate of cesarean sections as a measure of birth complications.[17]

In unadjusted regressions, the international framing has a positive but at conventional levels insignificant effect on financial contributions of respondents. Once adjusting for control variables, this coefficient turns significant at the 10% level. Respondents facing an international framing contribute on average more money in support of the SCC project than other midwives being confronted with the local framing. In the adjusted specification, their contribution is 1346 IDR higher. As we limit our sample to those respondents who were not exposed to the SCC prior to this experiment (see Chapter 4.2), we also check for robustness in the full sample. For this purpose, we estimate a regression, which controls for an interaction of the framing with the indicator for prior contact (see Table D4). Results are in line with those of the main regression.

5.2.1 | Channels: Previous exposure

In order to understand in more detail why respondents show stronger support toward projects implemented by international actors as compared to local implementers, we investigated *previous exposure* as a mechanism that is likely to influence the behavior of respondents. Previous exposure is one prominent factor shaping ideas and attitudes. Hence, it might play a role in whether respondents have been in contact with locally or internationally led projects in the past. Their respective experiences are likely to influence their present reactions to the intervention. Investigating the variation in exposure to international and local project implementers allows us to generate more general insights for locations with differing presence of the

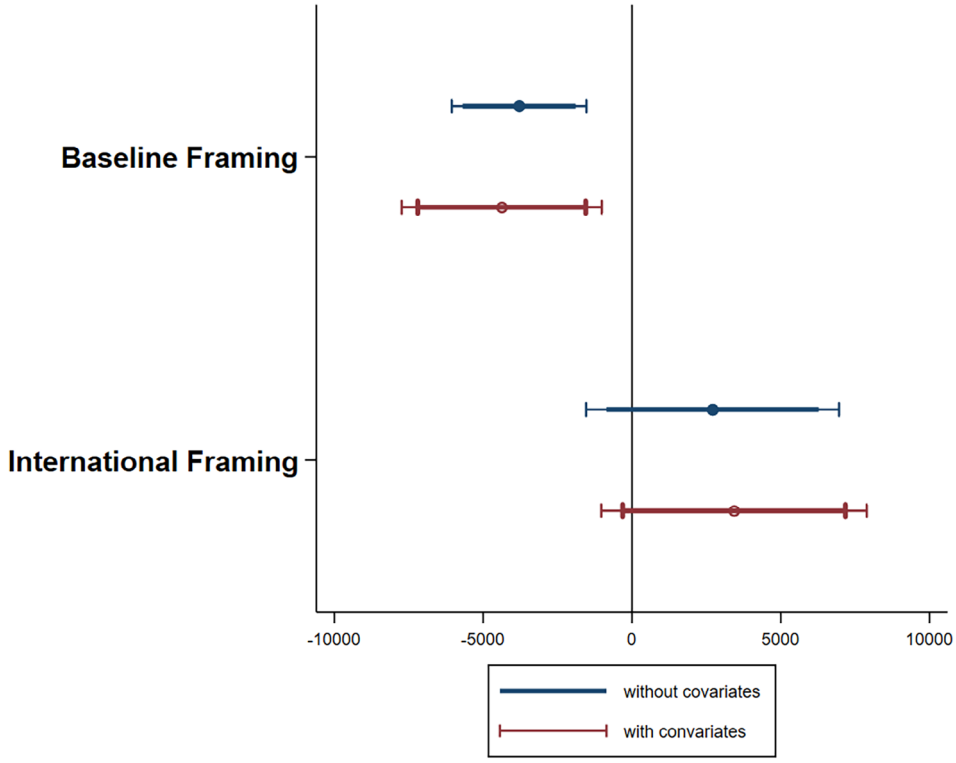

FIGURE 2 Framing experiment—Previous experience. This Figure depicts added-up coefficients for the interaction terms of previous participation in both local and international projects in the baseline (local) and the treatment (international) framing condition. Covariates correspond to Table 1 and include a variable indicating the facility type, a continuous variable of self-assessed resource access, as well as the facility-level cesarean section rate. Other interactions for participants, who either have participated in only local or international projects were included in the regression as well. The comparison group had no previous experience with either actor and faced a local framing. Errors are clustered at the facility level. The thick bars refer to the 10 percent and the thin bars to the 5 percent confidence interval. [Colour figure can be viewed at wileyonlinelibrary.com]

respective actors. Aceh is specific due to the activity of manifold—oftentimes international—donors in response to the human tragedy of the Indian Ocean tsunami.

Following Equation (2), Figure 2 displays the added-up point estimates as well as confidence intervals for the triple interactions. Those interactions consider our experimental framing with the binary variables indicating if respondents have already participated in international or local research projects in the baseline (local) and the treatment (international) framing condition (e.g., β_1*International Framing*participation$_{local}$*participation$_{international}$ + β_2*participation$_{local}$ *participation$_{international}$ + β_3*International Framing*participation$_{local}$ + β_4*participation$_{local}$ + β_5*International Framing*participation$_{international}$ + β_6*participation$_{international}$ + β_7* International Framing). For the respondents in the international framing condition, the point estimates are obtained by adding-up β_1 to β_7, whereas for the respondents in the local framing β_2,β_4, and β_6 are considered.[18] These point estimates correspond to those presented in Table D7 in the Appendix and try to ease interpretation. To achieve this, the different options were coded as categories and should be interpreted as the difference from the base category "No Experience with International Experts – No Experience with Local

Experts – No International Framing." Respondents, who have worked with both international and local actors are of particular interest due to the comparisons they can draw. As before, the framing indicator equals one for the international framing treatment and zero for the local framing treatment.

Blue bars in Figure 2 indicate the coefficients of regressions without covariates and red bars point estimates, which were adjusted for the unbalanced control variables. Regarding confidence intervals, thick bars refer to the 10% and thin bars to the 5% interval.

The Figure indicates a distinct pattern for health workers, who have been exposed both to an international and local project in the past. Our results indicate a lower contribution of 3800–4400 IDR (e.g., 0.22–0.28 US$) if those health workers face the local framing (p-value: .006 without control variables; p-value: .024 with control variables).[19] In contrast, this implies that the attitude toward the intervention is significantly more positive if respondents knowing both international and local actors are facing the international framing condition.

Thus, the results from Figure 2 suggest that the positive effects of the international framing are driven by previous experience with the respective implementer. The reduced willingness to contribute to local projects is most pronounced if respondents have participated in both local and international projects.

5.2.2 | Qualitative research

Qualitative data based on 66 surveys with health practitioners were collected to provide a clearer understanding of how experience contributes to a higher support of interventions perceived as international. Answers to the question "Please describe your experience working with international teams. What did you find surprising?" suggest that positive attitudes toward internationals are mostly linked to experiences of more structured implementation approaches (13 indications) and a higher perceived level of knowledge (4 indications). Moreover, in responding to the question "(W)hat are some of the strengths and challenges of international projects?" knowledge sharing (13 indications) and compliance with international standards (8 indications) were named as most important advantages. In line with a home bias argument (Fuchs & Gehring, 2017), health workers indicated language barriers as a relevant issue (3 indications).

This is in line with the positive and significant correlation of the international framing with positive perceptions of international control capabilities and skills of local implementers (Appendix Table D6).[20]

The additional qualitative evidence, thus, underlines that higher support for international projects is based on deeper perceptions of international/local implementation. These can, however, be highly context-specific, which will be discussed among other implications in the following section.

6 | DISCUSSION AND CONCLUSION

The success of a project depends critically on the support of the target population. This study investigates a potentially critical factor for supportive behavior evident in most development projects: the salience of local versus international agents in development projects. This is of

particular interest as the majority of interventions in the field of development economics are co-operations between local and international agents.

Using the introduction of a new tool, the SCC among health practitioners in Indonesia's Aceh province, we provide evidence from a real-world setting via a framing experiment. Results indicate that respondents are more supportive toward interventions (measured through monetary support) implemented by international actors as compared to solely locally led projects. This finding adds to previous studies, which suggest to increase support for the SCC specifically through an improved coaching (Delaney et al., 2017) or integration into the local context (Kumar et al., 2016). More generally, our study provides further support to previous research on behavioral reactions toward international and multilateral donor agencies (e.g., Milner et al., 2016; Winters et al., 2017). The striving for high visibility among international donors (Vollmer, 2012) and higher quality perceptions toward donors in comparison to the national government (Dietrich & Winters, 2015; Winters et al., 2017) are explanatory reasons for these results. Corresponding to this, our results indicate that health workers have a more positive perception of the skills of international agents.

Our results suggest that previous experience is pivotal. Those respondents that have already been exposed to previous internationally led interventions take a more positive stance toward future international projects. This relationship cannot be established for those who already participated in local research projects. In this respect, one has to consider that the experiment was conducted in a context in which previous exposure to international projects has been high and generally positive. The large exposure to various international as well as local actors in the aftermath of the Tsunami 2004 (Becerra et al., 2014; Doocy et al., 2007) facilitates the assessment toward the different implementers. However, this context of the ultimate human emergency, might have induced a more positive stance toward the international assistance and makes the interpretation specific to the context.[21] Further research in contexts with differential experiences with international actors could, thus, address external validity concerns (Lusk et al., 2006).

Overall, our results underscore the importance of responsible conduction of interventions by local as well as international agents as previous experience with the respective agents influences the attitude and support for future interventions. Generally, using framing as a tool to make a well-regarded implementing agent more salient might be a low-hanging fruit to increase supportive behavior of population groups in a cost-effective way (Bertrand et al., 2006). However, in order to increase the sustainability of projects, locally led implementation for fostering local ownership is critical. While the image of bilateral donors and international organizations may be particularly sticky (Nielson et al., 2019), positive strengthening local agents' capacities may improve experiences of the target group with locally led projects and prospects for scalability. Local actors, who implement development programs independently (or in equitable cooperation with internationals), could increase local ownership. This relates to the general debate on how aid can be delivered most successfully, and whether foreign funding undermines state legitimacy (e.g., Dietrich et al., 2018). Our research suggests that previous experiences with locally led projects matter for take-up by the target population. Thus, studying the effects of local capacities on target group's support appears as an important research endeavor. Here, using real-world settings would add to the applicability and transferability of research findings. Future studies may want to go one step further and manipulate the implementer identity in randomized controlled trials to measure the impact on intervention uptake (e.g., checklist usage).

ACKNOWLEDGEMENTS

We are especially thankful to Martin Bruder, Kerstin Grosch, York Hagmayer, Friederike Lenel, Holger Rau, Robert Schmidt, and Hendriek Yopin for their insightful feedback and support. We also thank Simone Dietrich, Axel Dreher, Victoria Gonsior, Marcela Ibañez, Ghida Karbala, Krisztina Kis-Katos, Stephan Klasen, Stephan Klingebiel, Christoph Kubitza, Sascha Kuhn, Dirk Landmann, Julia Leininger, David McKenzie, Anna Merkel, Rossa O·Keeffe-O·Donovan, Atika Pasha, Gert Pönitzsch, Sebastian Prediger, Jan Priebe, Miriam Romero, Santiago Saavedra, Sebastian Schneider, Armin von Schiller, Simone Schotte, Reinhard Weisser and colleagues from the chair in Heidelberg as well as the participants of seminars at Heidelberg University and the University of Göttingen and the members of the Research Training Group 1723 "Globalization and Development." Also, we are thankful for constructive feedback and comments from the anonymous reviewers. Moreover, we are grateful for excellent research assistance in Indonesia by Vivien Hülsen, Masyitah Masyitah, Cynthia Eka Putri, Teddy Thaher, Mutia Elvani, Hujjatul Balighah, and Lisa Rogge and for qualitative research support from Siobhan Doria and Farah Diba. Lastly, we sincerely thank our partners on the Safe Childbirth Checklist implementation in Indonesia Farah Diba, Marthoenis, Muhsin, Ichsan, Suryane Susanti and Samadi without whom this research could not have taken place. The study was facilitated by funding in the framework of Lower Saxonian Ministry of Science and Culture's "Reducing Poverty Risk" project, the German Research Foundation and the European Commission's Expert4Asia Program. Data from Indonesia was collected during Kuhnt's and Kaplan's employment with the University of Göttingen and Richert's employment with the University of Heidelberg. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

There are no conflicts of interest to be declared.

DATA AVAILABILITY STATEMENT

Data will be made available upon reasonable request.

ORCID

Lennart C. Kaplan https://orcid.org/0000-0003-3918-1606

ENDNOTES

[1] The general checklist was adapted to the country context and is available via the WHO Webpage, last accessed January 26, 2019.

[2] For a detailed description of the intervention, see the evaluation article of the main evaluation study (Kaplan et al., 2021).

[3] The framing effect became popular through its essential role in Kahneman & Tversky's, 1979 prospect theory in which they describe gambles either by their loss or gain probability. We consider an attribute framing, in distinction to risk or goal framings.

[4] Focusing on control facilities ensured that these midwives neither had yet received the SCC nor were in contact with the implementation team up to this point.

[5] We did purposefully not include a neutrally framed group in the framing experiment as development programs are always either conducted exclusively locally or have an international component. We believe that it is very unlikely that the implementer's identity is unknown to program participants, although salience might differ.

[6] This survey was included in the endline survey of the larger SCC intervention.

[7] All respondents chose to continue and participated in the following framing experiment.

[8] We focus on the traditionally employed monetary outcome as due to the costs incurred by the respondent this is likely to be the strongest measure. Estimates unB the additional outcomes provide qualitatively similar results and are available upon request.

[9] If respondents wanted to contribute, we offered them five options from 5000 to 25,000 IDR (equivalent to 0.4–1.9 US$) due to pragmatic reasons of specific top-up values.

[10] For the correlations, please see Table D6.

[11] Despite the medium sample size some control variables are unbalanced, as indicated in our balance test in Table D1.

[12] One needs to be aware that, especially, with a limited sample size omitted variables might not be homogenously distributed and, hence, it is not inherently clear, which other factors are correlated with our interaction variable of interest. However, balancing tests provided in Tables D1 and D2 underscore that previous participation is balanced across both framing treatments.

[13] Although the respective facilities were not exposed to the SCC, reasons for previous exposure might be a second job at another (treatment) facility (11.11% of respondents have a second job) or communication with other health practitioners within the district. Contact to midwives from other facilities is in this regard also significantly correlated with prior checklist contact.

[14] As a robustness check, we also report the full sample results, controlling for an interaction of prior contact with the treatment in Table D4.

[15] Due to four respondents that refrained from answering on control questions, the sample is reduced to $n = 167$ in our main specifications.

[16] Regressions on the alternative outcome measures yield qualitatively similar results and are available upon request.

[17] As a conservative robustness check, we also present random inference-based p values. Randomization inference takes the randomization explicitly into account and follows R.A. Fisher's idea of statistical inference via permutation tests of treatment allocation (Young, 2017). The idea is to assume uncertainty about the treatment allocation and compare the actual treatment allocation to possible alternative allocations.

[18] Confidence intervals were obtained via Stata's lincom command.

[19] Although this amount seems small, it corresponds to a small meal or 15 min of work of a health worker in the local context.

[20] We asked health practitioners if they would attribute certain characteristics rather to local or international researchers (e.g., skills, corruption, and financial capabilities) in order to carve out how those channels might affect support for the intervention. Those questions were asked intentionally after collecting the outcomes in order to not confound the results. However, this comes with the risk of justification bias. In fact, we find significant framing effects in our results, which are available upon request. Hence, we did not use those channels for further analysis. Yet, they might be still informative in terms of general attribute ascription.

[21] Despite the individual tragedies, parts of the population perceived the natural disaster as a chance to restart, as the successful reconstruction efforts coincided with the cessation of the Aceh insurgency after almost 30 years of combat (De Alwis & Noy, 2019). Moreover, Aceh might be specific due to its strong Muslim heritage and introduction of Islamic law in 2006.

[22] The framing effect became popular through its essential role in Kahneman and Tversky's (1979) prospect theory in which they describe gambles either by their loss or gain probability. There are three different types of framing approaches that have been described and used in the literature: the risky choice framing, goal framing and attribute framing, where we rely on the latter which makes certain characteristics of a choice or good more salient. Since then, framing experiments have been extensively applied in medical sciences both in hypothetical (Wilson et al., 1987) and real contexts, often related to message framing experiments, for example, with regard to smoking cessation, HIV screening as well as skin and breast cancer prevention (Detweiler et al., 1999; Kalichman & Coley, 1995; Schneider et al., 2001; Toll et al., 2007).

[23] As it is likely that respondents equate an international actor to a donor, we specifically addressed the relevant actors as researchers and professionals in our framing component.

[24] For this purpose, we talked to healthcare providers from different facilities, which were not part of the sampled institutions. In the Acehnese setting "local" is understood as "Acehnese" identity, whereby "Indonesian" would be an external concept. Certainly, it would have been of large interest to examine the difference between Acehnese and Indonesian implementers. However, due to power constraints, we decided to focus on this more specific framing without splitting the group and reducing the sample. The distinctness of "Acehnese" and "Indonesian" is also underlined by the fact that a small set of respondents named Indonesia and certain provinces as international countries. To deepen our understanding of the term "international" in the Acehnese context, we asked respondents to name the three countries, they first think of when hearing this term (see Figure C1 in Appendix). There is a large consensus among respondents regarding the main countries associated with "international," namely Germany (24%), Malaysia (19%), USA (13%), and Australia (8%). The high prominence of Germany among the foreign countries named, could first—of course—be attributed to the fact that parts of the implementing researchers, were German. Second, it is likely that Germany is indeed particularly present to the Acehnese people as it was the largest European donor after 2004s Tsunami (BBC, 2005). Moreover, Germany's reconstruction efforts were characterized by a strong focus on health interventions (German Federal Ministry for Economic Cooperation and Development (BMZ), 2005).

[25] After the debriefing, we offered participants to change their monetary contribution. Thirty-nine (16.5%) participants made use of this option. Generally, this led to an increase in contributions by on average one category (about 4200 IDR), but the amount is not contingent on the framing applied. The main analysis focuses on the pre-debriefing contribution, as we are interested in the framing effect.

[26] The Indonesian version of the experimental protocol is available upon request.

REFERENCES

Ashraf, N., Field, E., & Lee, J. (2014). Household bargaining and excess fertility: An experimental study in Zambia. *American Economic Review*, *104*(7), 2210–2237.

Banerjee, A., Duflo, E., Glennerster, R., & Kothari, D. (2010). Improving immunisation coverage in rural India: Clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *BMJ*, *340*, 1–9.

Banuri, S., Dercon, S., & Gauri, V. (2019). Biased policy professionals. *The World Bank Economic Review*, *33*(2), 310–327.

BBC. (2005). Germany takes lead in aid effort. Retrieved from http://news.bbc.co.uk/2/hi/europe/4148719.stm

Becerra, O., Cavallo, E., & Noy, I. (2014). Foreign aid in the aftermath of large natural disasters. *Review of Development Economics*, *18*(3), 445–460.

Bertrand, M., Mullainathan, S., & Shafir, E. (2006). Behavioral economics and marketing in aid of decision making among the poor. *Journal of Public Policy & Marketing*, *25*(1), 8–23.

Borchard, A., Schwappach, D. L. B., Barbir, A., & Bezzola, P. (2012). A systematic review of the effectiveness, compliance, and critical factors for implementation of safety checklists in surgery. *Annals of Surgery*, *256*(6), 925–933.

Bowles, S., & Polania-Reyes, S. (2012). Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature*, *50*(2), 368–425.

Bun, M. J., & Harrison, T. D. (2018). OLS and IV estimation of regression models including endogenous interaction terms. *Econometric Reviews*, *113*(522), 1–14.

Burgess, D. J. (2010). Are providers more likely to contribute to healthcare disparities under high levels of cognitive load? How features of the healthcare setting may lead to biases in medical decision making. *Medical Decision Making*, *30*(2), 246–257.

Burnside, C., & Dollar, D. (2000). Aid, policies, and growth. *American Economic Review*, *90*(4), 847–868.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, *90*(3), 414–427.

Cilliers, J., Dube, O., & Siddiqi, B. (2015). The white-man effect: How foreigner presence affects behavior in experiments. *Journal of Economic Behavior and Organization*, *118*, 397–414.

Cole, S., Giné, X., Tobacman, J., Topalova, P., Townsend, R. M., & Vickery, J. (2013). Barriers to household risk management: Evidence from India. *American Economic Journal: Applied Economics*, *5*(1), 104–135.

De Alwis, D., & Noy, I. (2019). Sri Lankan households a decade after the Indian Ocean tsunami. *Review of Development Economics*, *23*(2), 1000–1026.

Deck, C., & Jahedi, S. (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review*, *78*, 97–119.

Delaney, M. M., Maji, P., Kalita, T., Kara, N., Rana, D., Kumar, K., Masoinneuve, J., Cousens, S., Gawande, A. A., Kumar, V., Kodkany, B., Sharma, N., Saurastri, R., Singh, V. P., Hirschhorn, L. R., Semrau, K. E. A., & Firestone, R. (2017). Improving adherence to essential birth practices using the who safe childbirth checklist with peer coaching: Experience from 60 public health facilities in Uttar Pradesh, India. *Global Health: Science and Practice*, *5*(2), 217–231.

Detweiler, J. B., Bedell, B. T., Salovey, P., Pronin, E., & Rothman, A. J. (1999). Message framing and sunscreen use: Gain-framed messages motivate beach-goers. *Health Psychology*, *18*(2), 189–196.

Diba, F., Ichsan, I., Muhsin, M., Marthoenis, M., Sofyan, H., Andalas, M., Monfared, I., Richert, K., Kaplan, L. C., Rogge, L., Doria, S., Samadi, S., & Vollmer, S. (2019). Healthcare providers' perception of the referral system in maternal care facilities in Aceh, Indonesia: A cross-sectional study. *BMJ Open*, *9*(12). https://doi.org/10.1136/bmjopen-2019-031484

Dietrich, S., Mahmud, M., & Winters, M. S. (2018). Foreign aid, foreign policy, and domestic government legitimacy: Experimental evidence from Bangladesh. *The Journal of Politics*, *1*(80), 133–148.

Dietrich, S., & Winters, M. S. (2015). Foreign aid and government legitimacy. *Journal of Experimental Political Science*, *2*(2), 164–171.

Doocy, S., Rofi, A., Moodie, C., Spring, E., Bradley, S., Burnham, G., & Robinson, C. (2007). Tsunami mortality in Aceh province, Indonesia. *Bulletin of the World Health Organization*, *85*, 273–278.

Duflo, E., Glennerster, R., & Kremer, M. (2007). Chapter 61 using randomization in development economics research: A toolkit. *Handbook of Development Economics*, *4*(7), 3895–3962.

Easterly, B. W., Levine, R., & Roodman, D. (2004). Aid, policies, and growth: Comment. *American Economic Review*, *94*(3), 774–780.

Findley, M. G., Harris, A. S., Milner, H. V., & Nielson, D. L. (2017). Who controls foreign aid? Elite versus public perceptions of donor influence in aid-dependent Uganda. *International Organization*, *71*, 1–31.

Fuchs, A., & Gehring, K. (2017). The home bias in sovereign ratings. *Journal of the European Economic Association*, *15*(6), 1386–1423.

German Federal Ministry for Economic Cooperation and Development (BMZ). (2005). Review of German Government Funding for the Tsunami Relief Operations.

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don·t) work to modify behavior. *Journal of Economic Perspectives*, *25*(4), 191–210.

Grigoli, F., & Kapsoli, J. (2018). Waste not, want not: The efficiency of health expenditure in emerging and developing economies. *Review of Development Economics*, *22*(1), 384–403.

Haugen, A. S., Søfteland, E., Almeland, S. K., Sevdalis, N., Vonen, B., Eide, G. E., Nortvedt, M. W., & Harthug, S. (2015). Effect of the World Health Organization checklist on patient outcomes: A stepped wedge cluster randomized controlled trial. *Annals of Surgery*, *261*(5), 821–828.

Hess, S. (2017). Randomization inference with Stata: A guide and software. *Stata Journal*, *17*(3), 630–651.

Hong, F., Hossain, T., & List, J. A. (2015). Framing manipulations in contests: A natural field experiment. *Journal of Economic Behavior & Organization*, *118*, 372–382.

Hossain, T., & List, J. A. (2012). The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science*, *58*(12), 2151–2167.

Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, *302*(5649), 1338–1339.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, *93*(5), 1449–1475.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Kalichman, S. C., & Coley, B. (1995). Context framing to enhance HIV-antibody-testing messages targeted to African American women. *Health Psychology*, *14*(3), 247–254.

Kaplan, L. C., Ichsan, I., Diba, F., Marthoenis, M., Muhsin, M., Samadi, S., Richert, K., Susanti, S. S., Sofyan, H., & Vollmer, S. (2021). Effects of the world health organization safe childbirth checklist on quality of care and birth outcomes in Aceh, Indonesia: A cluster-randomized clinical trial. *JAMA Network Open*, *4*(12), e2137168.

Kumar, S., Yadav, V., Balasubramaniam, S., Jain, Y., Joshi, C. S., Saran, K., & Sood, B. (2016). Effectiveness of the WHO SCC on improving adherence to essential practices during childbirth, in resource constrained settings. *BMC Pregnancy and Childbirth*, *16*(1), 1–11.

Lichand, G., & Mani, A. (2016). Cognitive droughts. CAGE Online Working Paper Series 298.

Lusk, J. L., Brown, J., Mark, T., Proseku, I., Thompson, R., & Welsh, J. (2006). Consumer behavior, public policy, and country-of-origin labeling. *Applied Economic Perspectives and Policy*, *28*(2), 284–292.

Milner, H. V., Nielson, D. L., & Findley, M. G. (2016). Citizen preferences and public goods: Comparing preferences for foreign aid and government programs in Uganda. *The Review of International Organizations*, *11*(2), 219–245.

Nielson, D. L., Hyde, S. D., & Kelley, J. (2019). The elusive sources of legitimacy beliefs: Civil society views of international election observers. *The Review of International Organizations*, *14*(4), 685–715.

Nizalova, O. Y., & Murtazashvili, I. (2016). Exogenous treatment and endogenous factors: Vanishing of omitted variable bias on the interaction term. *Journal of Econometric Methods*, *5*(1), 71–77.

Payne, J. W., Sagara, N., Shu, S. B., Appelt, K. C., & Johnson, E. J. (2013). Life expectancy as a constructed belief: Evidence of a live-to or die-by framing effect. *Journal of Risk and Uncertainty*, *46*(1), 27–50.

Schneider, T. R., Salovey, P., Apanovitch, A. M., Pizarro, J., McCarthy, D., Zullo, J., & Rothman, A. J. (2001). The effects of message framing and ethnic targeting on mammography use among low-income women. *Health Psychology*, *20*(4), 256–266.

Toll, B. A., O·Malley, S. S., Katulak, N. A., Wu, R., Dubin, J. A., Latimer, A., Meandzija, B., George, T. P., Jatlow, P., Cooney, J. L., & Salovey, P. (2007). Comparing gain-and loss-framed messages for smoking cessation with sustained-release bupropion: A randomized controlled trial. *Psychology of Addictive Behaviors*, *21*(4), 534–544.

Tversky, A., & Kahneman, D. (1980). *The framing of decisions and the rationality of choice [Technical report]*. Stanford University of California, Department of Psychology.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

Viceisza, A. C. (2015). Creating a lab in the field: Economics experiments for policymaking. *Journal of Economic Surveys*, *30*(5), 835–854.

Vollmer, F. (2012). Increasing the visibility and effectiveness of development cooperation: How to reconcile two competing objectives? Technical Report 67.

WHO. (2018). Newborns: Reducing mortality. Retrieved from http://www.who.int/news-room/fact-sheets/detail/newborns-reducing-mortality

Wilson, D. K., Kaplan, R. M., & Schneiderman, L. J. (1987). Framing of decisions and selections of alternatives in health care. *Social Behaviour*, *2*(1), 51–59.

Winters, M. S., Dietrich, S., & Mahmud, M. (2017). Perceptions of foreign aid project quality in Bangladesh. *Research and Politics*, *4*(4), 1–6.

Workman, M., Lesser, M. F., & Kim, J. (2007). An exploratory study of cognitive load in diagnosing patient conditions. *International Journal for Quality in Health Care*, *19*(3), 127–133.

Young, A. (2017). Channelling Fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. Retrieved from http://personal.lse.ac.uk/YoungA/ChannellingFisher.pdf

How to cite this article: Kaplan, L. C., Kuhnt, J., & Richert, K. (2023). International or national implementers—Who is better? Evidence from a framing experiment. *Review of Development Economics*, 1–27. https://doi.org/10.1111/rode.13008

APPENDIX A: EXPERIMENTAL APPENDIX

In the aggregate, our experiment compares whether the salience of international versus local program implementers affects support for the respective project. Stressing certain aspects of a

particular situation among otherwise equivalent descriptions can lead to very different perceptions and behavioral reactions (Hossain & List, 2012; Johnson & Goldstein, 2003; Kahneman, 2003; Payne et al., 2013; Tversky & Kahneman, 1981). The result is what is called the *framing effect*.[22] Stressing certain aspects invokes different associations and leads to different evaluations by the decision-maker. Framing effects have been incorporated into theories on human behavior to explain deviations from rational choices (e.g., prospect theory). Their application to real-world decision-making can have important practical implications.

Our framing information reads as follows:

> "Among other researchers, [INTERNATIONAL/LOCAL] researchers took an active role in introducing the checklist to 17 facilities in Aceh province. The research team received approval from the provincial health office of Aceh. However, no funding was provided by the provincial health office. [LOCAL/INTERNATIONAL] research assistants and [INTERNATIONAL/LOCAL] health professionals with a lot of experience in delivery services were important partners and greatly supported the project."

In order to abstract from the specific actors within our stetting, we named different actors (e.g., researchers, practitioners).[23] A qualitative investigation was conducted prior to the experiment to ensure that the correct terms were used to describe "local" versus "international" agents.[24] To prevent potential effects through assumptions on political involvement, we specifically address the role of the provincial health office in the information given to the study participants. Further, to counter potential bias through speculations on the financial capabilities of different actors, we stress that funding of the intervention is ensured irrespective of the framing given to the participant.

After the experiment, all participants received a debriefing.[25] To create transparency on the use of the collected funds, we publicly made information on total amounts available after the end of the study and informed the participant about this procedure.

APPENDIX B: EXPERIMENTAL PROTOCOL

General remarks

If respondent asks you something, kindly answer by mentioning that you are only involved as an enumerator in the project and that you do not have any information on the Safe Childbirth Checklist.[26] Furthermore, please connect the respondent with the contact number, which has been stated before. Of course, if there are misunderstandings, you should repeat the provided information. However, please do not explain the information in different words.

Part A: "Now, we would like to present you a new tool and would like to learn about your opinion toward it." [*Before the start of the experiment (after the completed survey); give the 25,000 IDR voucher to the respondent*] "This is in appreciation of your time. Thank you very much. Subsequently, we will provide you with some information on a new tool for health care in Aceh province. After this, you can decide whether you want to take the money for yourself or if you want to contribute some for the implementation of this tool."

Part B: [*Enumerator: Please, read this introduction out aloud and clear.*] "During complex events, like performing a surgery or a delivery, people can be forgetful or might be distracted by

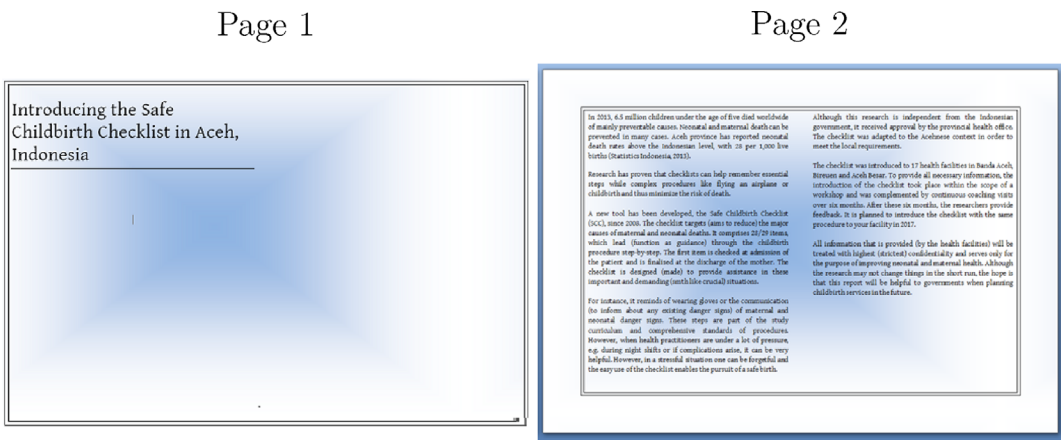

FIGURE B1 SCC leaflet. *Source*: Authors' own depiction. [Colour figure can be viewed at wileyonlinelibrary.com]

other emergencies or duties. This can potentially have terrible consequences, in the worst case losing the patient. Research proofs that checklists can save lives and prevent these mistakes. Like a surgeon is responsible for patients' lives in the operation theater, the delivery team can have great impact on the safety of mothers and babies. We would like to present you a new tool, which was developed especially for your everyday work: The Safe Childbirth Checklist. It comprises 30 easy-to-use items. The checklist begins with the admission of the patient and ends with the discharge of mother and baby from the hospital. In each delivery, the doctor or midwife fills in one checklist for every patient. You will fill in the checklist step by step and the checklist will remind you to perform the important steps during delivery. If you would like to know more about the checklist, here it is." [*Enumerator: Please hand a checklist copy over to the doctor or midwife.*] "For example, the checklist reminds you to perform easy things, which are nevertheless very important like hand washing." [*Enumerator: Show item "Confirm supplies are available to clean hands and wear gloves for each vaginal exam." on checklist*] "The checklist also reminds you to share important information with patients, including danger signs." [*Enumerator: Show item* "*Danger Signs*" *on checklist to the midwife or doctor*] "All these steps are already part of the study curriculum. Hence, every checklist item is easy to understand. Generally, most of the health workers already practice these important steps in the delivery process. The checklist just has the purpose to remind you of all the important steps during the delivery process. Especially, when health practitioners are under a lot of pressure, for example, during night shifts or if complications arise, it can be very helpful. For instance, a research study has proven that during surgeries simple checklists can help to reduce death rates even by almost half."

Part C: "Among other researchers, [*INTERNATIONAL/LOCAL*] researchers took an active role in introducing the checklist to 17 facilities in Aceh province. The research team received approval from the provincial health office of Aceh. However, no funding was provided by the provincial health office. [*LOCAL/INTERNATIONAL*] research assistants and [*INTERNATIONAL/LOCAL*] health professionals with a lot of experience in delivery services were important partners and greatly supported the project."

Part D: "I will now read to you information about the funding of the Safe Childbirth study conducted by the [*INTERNATIONAL/LOCAL*] researchers. The following is a page of paper containing information on the checklist." [*Enumerator: Please hand over the SCC leaflet to the participant*] (Figure B1).

> "The funds for the study have been used to implement the Safe Childbirth Checklist in 17 health facilities in Aceh province during October 2016. Funds are still available to introduce the checklist to 16 further facilities. The budget is enough to provide the 17 health facilities over six months with checklist copies. Therefore, every delivery during these six months can be conducted with the checklist. After this survey ends, the first six months of the checklist implementation are also over. There will be no funds remaining to provide additional checklists to those 17 health facilities, where the checklist was already introduced before."

Part E: "The researchers are collecting funds to be able to provide checklist copies at those health facilities. Are you willing to support the activity? Remember that the money collected will exclusively be used to provide checklist copies to the health facilities. The total amount of money that was contributed by all donors together will be made transparent. After finalizing the data collection, the amount of money collected will be published openly in every participating facility of this research. If you would like to support the activity, please decide on the amount of money you would like to contribute and note it down on the voucher. You can choose to not contribute at all, or you can give 5000, 10,000, 15,000, 20,000, or 25,000 IDR. Every contribution can help to conduct more deliveries with a Safe Childbirth Checklist. When you are done, please put the voucher in the envelope and seal it. If you do not wish to contribute anything, please put the number 0 on the voucher. In the end, only the aggregate amount of contributions from all participating facilities will be announced. Your individual contribution will be treated confidentially."

Part F: [*Enumerator: Read this introduction out aloud to the participant*] "During the following task you have to estimate the most chosen answer, which neither refers to the total amount nor the average. We have asked also other health practitioners/workers in the district how much is their willingness to contribute to the provision of checklist copies. Which amount do you think was contributed to the checklist copies by your colleagues per person at other facilities? This estimation is not at all related to your personal opinion. Instead, we would like you to estimate which amount of contribution that was given by most of the other health practitioners per person. For this question, if you assessed the most chosen amount per person correctly, you will be given an additional 10,000 IDR. If you estimated the right amount, the 10.000 IDR will be topped up to your phone credit together with the voucher within the next few days. The other health practitioners also had to choose to contribute 0, 5000, 10,000, 15,000, 20,000, or 25,000 IDR. Which category do you think was the most frequently chosen by the health workers?/Which amount do you think most other health workers chose to contribute per person?"

Part G: "Your facility is one of the other 16 facilities, where the research team would like to implement the Safe Childbirth Checklist. Experience shows that checklist use needs to be practiced with coaches regularly in order to make deliveries safer. How committed are you in investing your time to practice the use of the checklist in every week?"

Debriefing: "Thank you very much for your participation. We asked you previously several questions. The aim is to find out what is your opinion about *[local/international]* researchers and how this opinion influences your motivation to use the Safe Childbirth Checklist. The checklist was previously pilot tested in other countries around the world. This way the most crucial practices during child delivery were identified. The research collaboration was led by (international researchers) [...] and the World Health Organization. Local researchers [...] worked together with international researchers to adapt the checklist to the local context. Both parties hope that the Safe Childbirth Checklist can be implemented sustainably to serve as a tool for safe deliveries in Aceh province. If this information change your attitude toward

contributing to the checklist copies in any way, you are free to change your indicated contribution." [*Enumerator: If the respondent decides to change his/her contribution, please hand the envelope back.*]

APPENDIX C: FIGURES

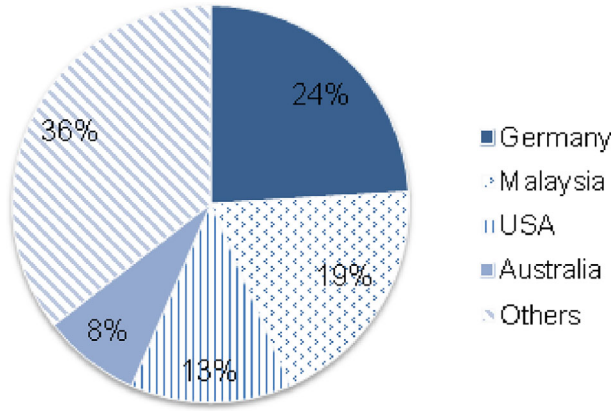

FIGURE C1 Distribution of "International" country perceptions. Based on "If you think of activities, programs or projects by internationals, which countries come first to your mind?". *Source*: Authors' depiction. [Colour figure can be viewed at wileyonlinelibrary.com]

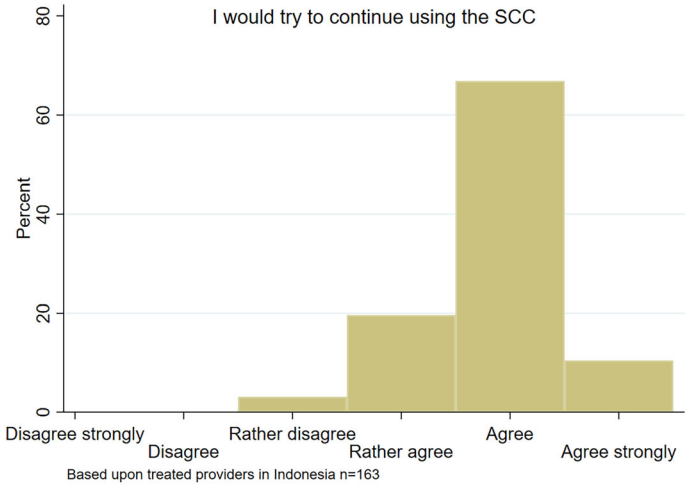

FIGURE C2 Intentions to use the Safe Childbirth Checklist—Indonesia. *Source*: Authors' calculation based on survey data. $N = 163$. [Colour figure can be viewed at wileyonlinelibrary.com]

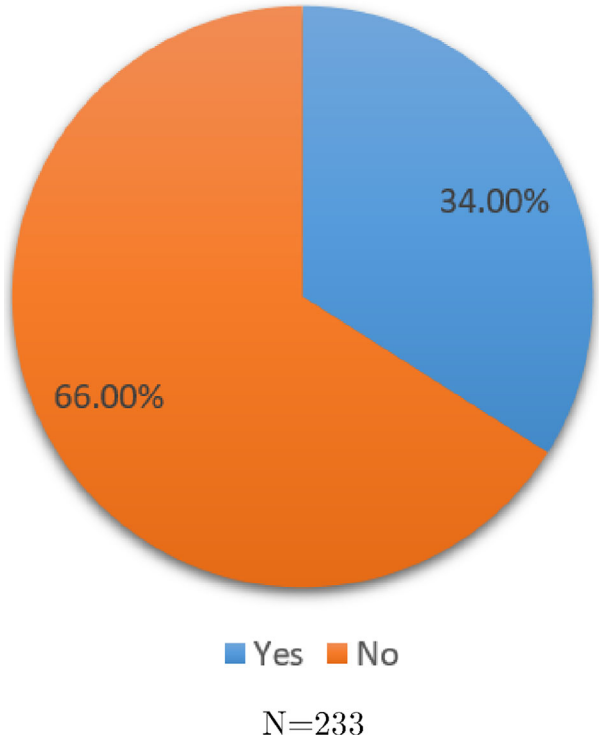

FIGURE C3 Actual use of the Safe Childbirth Checklist. $N = 233$. *Source*: Authors' calculation based on clinical observations. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE C1 Summary statistics for Indonesian data.

	N	Max	Min	Mean	SD
Actual behavior					
Active SCC use	219	1	0	0.389	0.489
Intended behavior					
Would try to use SCC even if copies not provided	163	6	3	4.847	0.634
Would recommend the SCC to fellow colleagues	163	6	2	5.092	0.495
Using the SCC in my professional role is	163	6	4	5.325	0.483
Ease to use SCC in work environment	163	6	4	5.141	0.565
SCC supported by superiors	163	6	4	5.828	0.439
Urban (1)—Rural (2)	163	2	1	1.515	0.501
CEmONC Service Provision 24/7	163	1	0	0.178	0.384
Facility Type: Community Health Center	163	1	0	0.589	0.494
Facility Type: Public Hospital	163	1	0	0.135	0.343
Facility Type: Private Hospital	163	1	0	0.190	0.394
Facility Type: Private Midwife Clinic	163	1	0	0.086	0.281
District: Aceh Besar	163	1	0	0.276	0.448
District: Banda Aceh	163	1	0	0.331	0.472
District: Bireuen	163	1	0	0.393	0.490

Descriptive statistics

Corresponding to the high pre-intervention commitment, which we observed among midwives, there is a high level of reported intentions (see Figure C2).

Yet, Figure C3 suggests a much lower level of actual uptake, which may point to a "know-do gap." Table C1 provides a more comprehensive overview of descriptive statistics corresponding to individuals' attitudes and facility characteristics.

APPENDIX D: ANALYTICAL APPENDIX

For the framing experiment, we find that the groups which were internationally or locally framed are balanced for the majority of variables (both in the full and reduced sample as depicted in Tables D1 and D2). Among the different observed variables, the minor differences pertaining to access to resources and facility type could be by chance. The average study participant was 33 years old (minimum: 21 years, maximum: 50 years), had 10 years of work experience (minimum: 0 years; maximum: 28 years), and 15 years of education (minimum: 12 years; maximum: 17 years).

Some of the respondents in the control group reported that they were previously in contact with the SCC. This does not imply a contamination of our control group per se, as the treatment was delivered on a clustered basis per facility in Indonesia. However, as there is informal exchange between healthcare personnel and shifts between facilities, midwives from other facilities might have heard about the checklist. Individuals with prior *contact to the checklist* might not have had *contact with the research team* and could, hence, still be receptive to the framing. First, including this group is more conservative as the framing should have a lower effect on the persons that are acquainted with the SCC and induce, thus, a downward bias. Second, individuals with prior contact to the checklist might react heterogeneously due to more comprehensive information. As a further robustness check, we estimate a regression in Table D4, which controls for an interaction of the framing with the indicator for past contact. Again, the positive and significant framing effect remains robust.

As the experimental outcome variables are all coded in a categorical (non-continuous) way, a probit regression model seems appropriate. Thus, we re-estimate the model in Table D5. The positive relationship between the framing and support for the intervention remains qualitatively unchanged. However, we prefer to present OLS estimates in the main part for ease of interpretation.

In order to understand the underlying pathways better, which explain the heterogeneous support for international and local actors, we also collected information on previous participation in local/international projects. This involves a trade-off: If prompting for those perceptions before framing individuals, reported support might be subject to justification of previously stated perceptions. If framing the respondents before collecting the perception measures, we might contaminate the latter data. We chose the second option to sustain the quality of our outcome measures. And indeed Table D6 indicates that the framing is significantly associated with several channel variables. For this reason, we prefer to rely only on previous project participation for our channel analysis. Although previous participation is self-reported, it is not perception based and, hence, less likely to be subject to justification bias. Table D6 supports this notion.

TABLE D1 Experimental balance—Full sample.

	Full *N*	Full Mean	Full SD	Control Mean	Control SD	Treat Mean	Treat SD	*p*-value difference
Puskesmas	234	0.650	–	0.716	–	0.56	–	.013**
Pub. Hosp.	234	0.162	–	0.134	–	0.2	–	.179
Priv. Hosp.	234	0.184	–	0.149	–	0.23	–	.116
Gender ($1 = m$, $2 = f$)	234	2.000	–	2.000	–	2.000	–	–
Age (years)	234	33.342	7.493	33.650	7.806	33.112	7.316	.593
Education (years)	234	15.047	0.527	15.020	0.603	15.067	0.462	.619
Experience (years)	234	9.603	7.271	9.690	7.736	9.537	6.979	.886
Sufficient income	234	3.209	1.008	3.160	1.012	3.246	1.014	.526
Financial problems	230	0.348	–	1.720	–	1.642	–	.081*
Strategic donation	213	4.432	1.264	4.710	1.225	4.627	1.296	.564
Social acc. Index	234	3.410	0.838	3.450	0.821	3.381	0.857	.513
Social acc. # 1	234	4.966	0.690	5.000	0.778	4.940	0.622	.480
Social acc. # 2	234	4.568	1.027	4.600	0.932	4.545	1.101	.650
Social acc. # 3	231	5.342	0.558	5.310	0.506	5.366	0.595	.172
Social acc. # 4	231	4.641	1.074	4.694	1.069	4.602	1.087	.475
Social acc. # 5	234	2.231	1.254	2.250	1.298	2.216	1.235	.784
Paperwork too much	234	2.808	1.343	3.000	1.497	2.664	1.195	.173
Routines ease work	234	5.167	0.734	5.150	0.626	5.179	0.764	.660
Previous SCC XP	234	0.278	0.450	0.240	0.429	0.306	.463	.267
Resource access	234	3.470	0.517	3.530	0.502	3.425	0.526	.080*
Team effic.	234	5.244	0.513	5.220	0.462	5.261	0.547	.570
Part.n loc. projects	234	0.167	–	0.130	–	0.194	–	.235
Part. int. projects	234	0.103	–	0.120	–	0.090	–	.511
Part. donor projects	234	0.094	–	0.080	–	0.104	–	.511

Note: Based upon the full sample with *N* denoting the number of observations, where 134 individuals were in the treatment and 100 in the control group. SD gives the standard deviations, which are not depicted for binary outcomes. Proportions in the two groups are significantly different from each other. Asterisks indicate *p* values based on standard errors clustered at the facility level: $*p < .1$, $**p < .05$, $*** p < .01$.

Point Estimates – Previous Experience: Table D7 displays the results for the interaction of our experimental framing with the binary variables indicating if respondents already participated in international or local research projects. The results in columns (1a and 1b) are structured to compare respondents with similar previous experience (participation in international/local projects) across framings. The corresponding comparison group are locally framed respondents, who did neither participate in a local nor in an international project. Row I and II show that if a person had been exposed both to an international and local research project in the past, their contribution is approx. 5400–6400 IDR (e.g., 0.35–0.41 US$) higher if framed internationally. Thus, the effect of the *attitude* toward the intervention in the unadjusted and adjusted specification is significantly higher if respondents know both implementers and are framed

TABLE D2 Experimental balance—Reduced sample.

	Full *N*	Full Mean	Full SD	Control Mean	Control SD	Treat Mean	Treat SD	*p*-value difference
Puskesmas	170	0.7	–	0.763	–	0.618	–	.145**
Pub. Hosp.	170	0.1	–	0.145	–	0.065	–	−.080*
Priv. Hosp.	170	0.2	–	0.172	–	0.237	–	−.065
Gender (1 = m, 2 = f)	170	2.000	–	2.000	–	2.000	–	–
Age (years)	170	32.359	6.997	33.118	7.680	31.774	6.395	.232
Education (years)	170	14.994	0.516	14.974	0.565	15.011	0.478	.742
Experience (years)	170	8.888	7.094	8.974	7.494	8.849	6.824	.908
Sufficient income	170	3.200	1.069	3.118	1.083	3.269	1.065	.348
Financial problems	170	1.741	–	1.763	–	1.720	–	.396
Strategic donation	170	4.606	1.411	4.658	1.381	4.581	1.440	.613
Social acc. index	170	3.329	0.827	3.316	0.852	3.344	0.814	.808
Social acc. # 1	170	5.000	0.738	4.987	0.887	5.011	0.599	.834
Social acc. # 2	170	4.459	1.142	4.461	1.026	4.462	1.239	.991
Social acc. # 3	170	5.429	0.584	5.408	0.521	5.452	0.634	.436
Social acc. # 4	167	4.545	1.063	4.649	1.065	4.457	1.063	.239
Social acc. # 5	170	2.118	1.286	2.184	1.334	2.065	1.258	.375
Paperwork: too much	170	2.906	1.364	3.145	1.547	2.720	1.174	.150
Routines ease work	170	5.100	0.727	5.079	0.648	5.151	0.722	.471
Access to resources	170	3.441	0.498	3.513	0.503	3.387	0.490	.060*
Team effic. indicator	170	5.200	0.443	5.158	0.434	5.226	0.445	.459
Part. in loc. projects	170	1.829	–	1.868	–	1.796	–	.131
Part. in int. projects	170	1.918	–	1.895	–	1.935	–	.272
Part. in donor projects	170	1.935	–	1.934	–	1.935	–	.959

Note: Based upon the reduced sample excluding observations with prior contact to the checklist. *N* denotes the number of observations, SD gives the standard deviation. Standard deviations are not depicted for binary outcomes. Proportions in the two groups are significantly different from each other. Asterisks indicate *p* values based on standard errors clustered at the facility level: *$p < .1$, **$p < .05$, *** $p < .01$.

internationally (*p*-value: .0077 and .0080, respectively). Moreover, if respondents who face the local framing were only exposed to international and not to local projects, they do contribute significantly less if locally framed, significant without adjusting for controls (*p*-value: .0000 and .1129, respectively). Those estimates suggest that the positive effects of the international framing are driven by previous experience with the respective implementer. The reduced willingness to contribute to local projects is most pronounced if respondents have participated both in local and international projects.

TABLE D3 Framing experiment—Wild bootstrapped SE.

Financial contribution in support of SCC project (in IDR)			
	(a)	**(b)**	**(c)**
Framing: 1 = "internat."	756	1346*	1394*
p-value	(.290)	(.065)	(.058)
WB *p*-value	(.268)	(.042)	(.045)
N	167	167	167
Control variables	No	Yes	Yes
Mean of dep. var.	4758	4758	4758
SD of dep. var.	4711	4711	4711

Note: Same control variables as in Table 1. Standard errors (SE) are clustered at the facility level and wild bootstrapped due to limited cluster number (13) for the specifications indicated as "WB *p* values," following Cameron et al. (2008). Asterisks indicate *p* values according to $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

TABLE D4 Framing experiment—Interaction with prior contact.

Financial contribution in support of SCC project (in IDR)			
	(a)	**(b)**	**(c)**
Framing: 1 = "internat."	755.7971 (0.289) [0.344]	1221* (0.076) [0.112]	1236* (0.073) [0.121]
Resource Access		−498 (0.417)	−366 (0.592)
Pub. Hospital		−3215** (0.021)	−3000* (0.051)
Priv. Hospital		−2101 (0.312)	−1650 (0.594)
Pustu		−2241* (0.051)	−2300** (0.042)
C − Sec. Rate			−14 (0.726)
Int. Framing × Prior Contact	−413 (0.772)	−290 (0.838)	−240 (0.864)
Prior Contact	364 (0.735)	1030 (0.299)	1046 (0.299)
N	230	230	230
Control variables	No	Yes	Yes
Mean of dep. var.	4758	4758	4758
SD of dep. var.	4711	4711	4711

Note: The base category is No Prior Contact and Local Framing. RI *p* values are computed with a permutation test based on Hess (2017). RI *p* values/*p* values in brackets/columns. Asterisks indicate *p* values based on standard errors clustered at the facility level: $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

TABLE D5 Framing experiment—Ordered probit results.

Financial contribution in support of SCC project (categorical)			
	(a)	**(b)**	**(c)**
Framing: 1 = "internat."	756	1346*	1394*
p-value	(.290)	(.065)	(.058)
WB *p*-value	(.268)	(.042)	(.045)
N	167	167	167
Control variables	No	Yes	Yes
Mean of dep. var.	4758	4758	4758
SD of dep. var.	4711	4711	4711

Note: See Table 1. Reported coefficients are not transformed and represent ordered probit coefficients. Standard errors (SE) are clustered at the facility level. Asterisks indicate *p* values according to $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

TABLE D6 Framing experiment—Association with potential channel variables.

	Control capabilities	Implementation skills	Funding capabilities	Accountability	Trust foreign countries	Participation int. project	Participation loc. project
Framing: 1 = "internat."	0.802***	0.774***	0.604***	0.445*	0.045	0.023	−0.065
SE	0.214	0.210	0.188	0.243	0.051	0.047	0.055
p-value	.002	.003	.007	.090	.393	.638	.257
WB *p*-value	.004	.008	.008	.118	.374	.719	.224
N	230	230	230	230	230	230	230

Note: All specifications are based on the full sample. All specifications include a variable indicating the facility type, a binary variable indicating if the respondent had financial problems, a composite index of social desirability variables, and a variable indicating the subjective perception of the amount of paperwork. Standard errors (SE) are clustered at the facility level. We present results based on clustered SE indicated as "*p* values" and wild bootstrapped due to limited cluster number (13) for the specifications indicated as "WB *p* values," following Cameron et al. (2008). Asterisks indicate *p* values according to: *$p < .1$, **$p < .05$, *** $p < .01$.

TABLE D7 Framing experiment—Previous experience (point estimates).

Outcome: Financial contribution in support of SCC (in IDR)		
	(a)	**(b)**
(I.) International Framing (1) × Int. participation (1) × Loc. Participation (1)		
β	5362	6425
p-value	.217	.215
(II.) International Framing (0) × Int. participation (1) × Loc. Participation (1)		
β	1149	−1001
p-value	.575	.751
Coefficient Equality Row (I) & (II)	0.0077	0.0080
(III.) International Framing (1) × Int. participation (0) × Loc. Participation (1)		
β	−3029	−1866
p-value	.145	.303
(IV.) International Framing (0) × Int. participation (0) × Loc. Participation (1)		
β	−149 (0.918)	6 (0.997)
Coefficient Equality Row (III) & (IV)	0.4751	0.8195
(V.) International Framing (1) × Int. participation (1) × Loc. Participation (0)		
β	3280	1921
p-value	.114	.430
(IV.) International Framing (0) × Int. participation (1) × Loc. Participation (0)		
β	−4792***	−3380
p-value	.000	.186
Coefficient Equality Row (V) & (VI)	0.0000	0.1129
(VII.) International Framing (1) × Int. participation (0) × Loc. Participation (0)		
β	886	1325
p-value	.344	.153
N	165	165
Control variables	No	Yes

Note: Based on triple interactions of the framing (where International Framing (0) refers to the baseline condition) and binary indicators of previous experience with the respective actors. Standard errors (SE) are clustered at the facility level. Asterisks indicate p values according to *$p < .1$, **$p < .05$, *** $p < .01$.