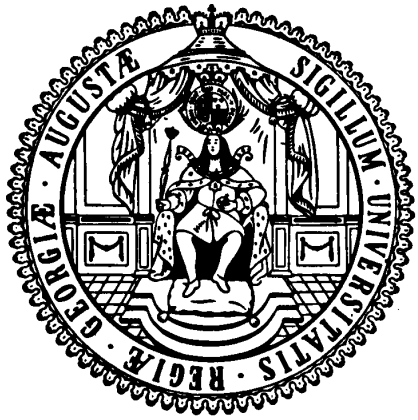
Courant Research Centre

'Poverty, Equity and Growth in Developing and Transition Countries: Statistical Methods and Empirical Analysis'

Georg-August-Universität Göttingen
(founded in 1737)



Discussion Papers

No. 261

What Explains the Uptake of Development Interventions?

Lennart C. Kaplan, Jana Kuhnt, Katharina Richert, and Sebastian Vollmer

April 2019

Platz der Göttinger Sieben 5 · 37073 Goettingen · Germany
Phone: +49-(0)551-3921660 · Fax: +49-(0)551-3914059

Email: crc-peg@uni-goettingen.de Web: http://www.uni-goettingen.de/crc-peg

What Explains the Uptake of Development Interventions?

Considering the Theory of Planned Behaviour*

Lennart C. Kaplan[a], Jana Kuhnt[a], Katharina Richert[b,c], and Sebastian Vollmer[d]

[a] *German Development Institute*
[b] *Center for Evaluation and Development*
[c] *University of Mannheim*
[d] *University of Göttingen*

April 30, 2019

*We thank Holger Rau, York Hagmayer, Kerstin Grosch, Ghida Karbala, Christoph Kubitza, Hendriek Yopin, Friederike Lenel, Robert Schmidt, Jan Priebe, Anna Merkel, Martin Bruder, Simone Dietrich, Sascha Kuhn, Dirk Landmann, Julia Leininger, David McKenzie, Sebastian Schneider, Miriam Romero, Axel Dreher, Stephan Klasen, Marcela Ibañez, Sebastian Prediger, Gert Pönitzsch, Atika Pasha, Rossa O'Keeffe-O'Donovan, Santiago Saavedra, Reinhard Weisser, Stephan Klingebiel, Victoria Gonsior, Armin von Schiller 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" for valuable feedback on the design and implementation of the field experiment conducted in Indonesia. Moreover, we are grateful for excellent research assistance 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, Samadi) and Pakistan (Ashfa Hashmi), without whom this research could not have taken place. The study was facilitated by funding in the framework of GIZ's "Reproductive, Maternal and Newborn Health" Project, Lower Saxonian Ministry of Science and Culture's "Reducing Poverty Risk" project, the German Research Foundation and the European Comission's Expert4Asia Program. Data from Indonesia and Pakistan were collected during Kuhnt's, Kaplan's and Vollmer's employment with the University of Göttingen and Richert's employment with the University of Heidelberg.

Abstract

A crucial prerequisite for the success of development interventions is their uptake in the targeted population. We use the setup of an intervention conducted in Indonesia and Pakistan to investigate dis-/incentivizing factors for program's uptake and support. Making use of a framework grounded in psychological theory, "The Theory of Planned Behaviour," we consider three determinants for intervention uptake: personal attitudes, the social influence of important others and the perceived ease of intervention use. As most development interventions are characterized by a cooperation among local and international agents, we investigate further a potentially important dis-/incentivizing factor: the salience of the implementer's background.

Our findings show that attitudes, important others and ease of intervention use are indeed associated with increased uptake in our two culturally different settings. Conducting a framed field experiment in Indonesia we show further 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. To strengthen supportive behaviour by the target population for locally led projects, it is essential to strengthen local capabilities to create positive experiences.

Hence, our results encourage development research and cooperation, first, to consider personal attitudes, the social influence of important others and the perceived ease of intervention use in the design of interventions in order to increase uptake. Second, depending on the country context, implementers should consider the previous experience with and attitude towards partners – either local or international – when aiming to achieve behavioural change.

Keywords: Theory of Planned Behaviour; Framed Field Experiment; Implementation Research; Public Health

1 Introduction

A large focus in the literature studying development cooperation naturally lies on its effectiveness. On the macroeconomic cross-country level, the effectiveness of aid is studied to an impressive extent, while results are still inconclusive (Burnside and Dollar, 2000; Easterly et al., 2004). In focus of the literature typically stand donor (Berthélemy, 2006; Minasyan et al., 2017) and recipient characteristics (e.g., Dollar and Pritchett, 1998; Rajan and Subramanian, 2008). Much less attention is drawn to the specific implementation features of development interventions, which might likewise and very likely predict interventions' success. Take for instance two very similar interventions on HIV/Aids education for young people in Uganda from Kinsman et al. (2001) and Karim et al. (2009). While Karim et al. (2009) show quite positive effects of the intervention on female participants with regard to increased condom use, Kinsman et al. (2001) see almost no effect of their large-scale intervention. Can we accordingly assume that HIV/Aids education works in all evaluated eight districts, but Masaka, where Kinsman et al. (2001) conducted their study? Alternatively in 2009, but not in 2001? Possible, but unlikely. The probability is higher that the implementation strategy, which Karim et al. (2009) tested, was more successful in achieving behavioural change than the approach evaluated by Kinsman et al. (2001) in the given setting. Uptake by the target population is one of the major factors influencing the success of an intervention. However, what influences the uptake of a development intervention? And under what circumstances is the target population more likely to support the program? At the heart of development interventions is regularly the aim to change human behaviour – generally as a mediator to reach a certain goal (e.g., increased use of condoms to reduce sexually transmitted diseases). Limited participation or support from the respective target population challenges these interventions (e.g., Banerjee

et al., 2010; Cole et al., 2013). In this study, we want to address the puzzle of uptake of and support for development interventions and examine "dis-/incentivizing factors." A systematic and deep understanding of what drives behavioural change in response to development programs is in high demand and is partly acknowledged by building a theory of change before implementation (Nayiga et al., 2014; Rogers, 2014). However, the application of a general framework is missing (Duflo et al., 2007; World Bank, 2015). Still most interventions analyzed in the field of development economics predominantly rely on monetary incentives to increase uptake. Other important drivers of human behaviour have attracted limited attention (**?**). This is the case, despite insights from behavioural economics stressing the importance of non-monetary incentives that shape human motivation and behaviour (e.g., Gneezy et al., 2011; Bowles and Polania-Reyes, 2012), and scholarly work showing that these factors play a role in the successful design of interventions (e.g., Banerjee et al., 2010; Cole et al., 2013; Ashraf et al., 2014).[1]

Being confronted with low uptake rates in two comparable interventions conducted by ourselves in Pakistan and Indonesia, we investigate possible dis-/incentivizing factors that might help explain why some people are more engaged while others are not. In order to systematically analyze drivers for non-/supportive behaviour we make use of a psychological theory called the "Theory of Planned Behaviour" (TPB). While there exist other theories aiming at explaining behavioural patterns, the TPB is the most established one that has been applied to a variety of different contexts (Blue, 1995; Armitage and Conner, 2001). It provides a straightforward framework to identify and respond to facilitating and hindering factors related to human behaviour. To the best of our knowledge, however, the framework has not yet been used to explain behavioural response to interventions in the field of development economics.

[1]These factors "disturbing" the rational decision-making are acknowledged by economists (here often-called psychological biases and cognitive limitations) and insights from behavioural economics are increasingly applied to public policy (e.g., Behavioural Insights Team in the UK; Mind, Behavior and Development Unit at the World Bank; Madrian (2014)).

The TPB proposes three determinants that influence human behaviour: the individual's attitude towards the intervention, subjective norms, and the individual's sense of behavioural control. We investigate the potential relationship of these determinants with intended and actual uptake rates within the setting of two real-world interventions. More specifically, we consider the introduction of the World Health Organization (WHO)'s Safe Childbirth Checklist (SCC) in Pakistan's Khyber Pakhtunkhwa province (Kuhnt and Vollmer, 2018) and Indonesia's Aceh province (Diba et al., 2018). Evidently, the checklist can only be effective if health personnel complies with the intervention and actually uses the SCC. Hence, the behaviour in question is the uptake (use) of the checklist during deliveries. One parameter common to almost all development interventions is the nexus between local and international implementers. Also during our project we realized that the international or local association of the implementing agents is likely to influence the behaviour of the target group towards the project. This is in line with recent research where implementer's characteristics as a softer preconditions for the support of interventions are found to play a role (e.g., Cilliers et al., 2015; Findley et al., 2017). Accordingly, we deepen our analysis of this behavioural determinant by conducting an additional framed field experiment. Within the context of the Indonesian SCC intervention, we assess whether health personnel's support towards checklist use changes conditional on whether the participation of local or international agents in the project is highlighted.[2]

Our results show that intended and actual uptake of the SCC in both country settings are indeed positively related to all three TPB determinants. A more positive attitude towards the project, greater behavioural control as well as supportive subjective norms are all related to increased uptake of the SCC in Indonesia and Pakistan. Hence, we argue that the TPB can help disentangling the puzzle of heterogeneous

[2]For a visualization of our study design, see Figure 2.

engagement by the target group and can serve as a guideline in determining and shaping factors affecting intervention uptake. Focusing on the implementation design on stimulating these factors is, thus, likely to increase the success of interventions through increased support and consequently higher participation rates among the targeted population. Further, our framed field experiment indicates that change in support for the 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 behavioural reactions towards international research projects. Hence, in the Indonesian context it seems to have advantages to stress the international nature of programs over solely locally organized projects. However, to support local ownership and successful local project implementation our results stress the importance of strengthening local capabilities to create positive exposure to locally led projects.

The study is structured as follows: Section 2 describes the background of our study. Section 3 introduces the "dis-/incentivizing factors" and gives background on the "Theory of Planned Behaviour". Section 4 links this framework to our interventions and describes our research design and data. Section 5 elaborates on the methods used, and the results are described in Section 6. Section 7 discusses the generalizability and policy relevance of the results and concludes the study.

2 Background

This study systematically considers dis-/incentives shaping behaviour. The analysis considers a light-touch health intervention, where take-up rates are more likely to be predicted via behavioural factors than due to technical constraints. Motivated by low uptake by the target group, we consider potential dis-/incentivizing factors in two

local contexts: Indonesia and Pakistan. On the one hand, relying on two distinct samples enables us to increase external validity of the investigated factors, which is one main concern with field studies. On the other hand, the comparison contributes to an understanding of heterogenous effects of those "dis-/incentivizing factors."

Two-thirds of maternal and new-born deaths globally occur due to causes, which could largely be prevented if well-established essential practices were followed (WHO, 2018). The WHO Safe Childbirth Checklist (SCC) initiative aims at providing health personnel with a four-page checklist to be used around the delivery process. The checklist entails the essential practices addressing the major risk factors for mothers and children in low and middle income countries.[3] Experience from other medical fields suggests 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 behavioural economics suggest that human behaviour 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., Croskerry, 2002; Burgess, 2010; Hoffman et al., 2011; Deck and Jahedi, 2015; Lichand and 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., Workman et al., 2007; Borchard et al., 2012; Haugen et al., 2015). Our international research teams implemented the checklist in collaboration with local partners. We used a light-touch approach in both country settings, which are described subsequently.[4]

Indonesia has invested large resources to improve its health care culminating in

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

[4]For a detailed description of the interventions, see the evaluation articles of the main evaluation studies (Diba et al., 2018; Kuhnt and Vollmer, 2018).

the introduction of a national health insurance (Jaminan Kesehatan Nasional) in 2014. Those investments are linked to an increased number of births attended by skilled health care providers (e.g., midwives or doctors). We conducted our study in Aceh province, which was after 30 years of civil war and 2004's tsunami subject to massive reconstruction efforts of the national government and international donors. More specifically, funds were used to establish an infrastructure of well-equipped health facilities (community health centers as well as hospitals), which complements the system of village and private midwives. Using a clustered randomized control design, we evaluated the SCC in 16 of those health facilities, while 16 additional facilities served as a control group. We focused our assessment on health facilities (in contrast to individual providers) as we hypothesized that existing quality management systems as institutions would facilitate implementation. Moreover, seniority plays a predominant role in the Indonesian society. Thus, existing hierarchies in health facilities enabled us to use supervisors' engagement as leverage to motivate the staff. Engagement was supported via a motivating launch event informing health personnel about the checklist's benefits for their everyday work, complemented by eleven coaching visits over the following six months.

In Pakistan, the study was conducted in two districts, Haripur and Nowshera, of the province Khyber Pakhtunkhwa (KP) in the Northwest of Pakistan. Improvements in maternal and new-born health are high on the national policy agenda and were recently endorsed in the 'National Vision for Coordinated Priority Actions 2016-2025' (WHO, 2016). To reach this goal improving the quality of skilled providers (including facilities and individuals) is essential. In close cooperation and with the support of the local authorities, the SCC was evaluated among 17 health facilities (of different size) and 149

individual health care providers (community midwives and lady health visitors)[5] using a cluster randomized control design.[6] The mix of providers is representative of the public health system in the two districts. Particularly the individual providers are only loosely attached to the local government structures (e.g., through provision of medication and equipment) and de facto there is very little oversight of their activities. In order to ensure knowledge on the SCC and motivate uptake, we conducted standardized trainings on the checklist for all health staff and launched it via events in the health facilities. Engaging local political authorities in this process, we ensured their support, which was important for the cooperation of the larger health facilities. The intervention was complemented with on average one monthly monitoring visits by the local project coordinator.

While implementing the same tool and following similar implementation procedures, the respective context differs. This allows us to investigate the role of dis-/incentives for intervention uptake in a more heterogeneous manner and herewith establish a potentially greater external validity of our findings.

3 "Dis-/Incentivizing Factors for uptake and support"

Why did uptake not work perfectly? Despite a high commitment of health care providers during the launch events of the Safe Childbirth Checklist, uptake was lower than expected. The puzzle of this study is, thus, as follows: If health personnel *know*

[5]Community midwives in Pakistan are trained midwives, who operate on their own within local, often rural, communities. Lady Health Visitors are mid-level health care providers with a high-school diploma and a two-year medical training, providing health care to mothers and children under five years.

[6]According to the evaluation design the SCC was randomly implemented in roughly half of those providers.

that the checklist entails necessary essential practices supporting the safety of deliveries, why would they decide not to *use* the checklist. Following the ideas of the rational choice theory that describes independent agents striving to maximize their utility (Simon and Feldman, 1959), the deviation should be a matter of *incentives*, assuming the availability of information or technical knowledge and necessary equipment.[7] In order to understand factors that might have constrained the uptake, we, hence, continue with a more systematic overview of potentially important (dis-)incentivizing factors.

3.1 (Dis-)Incentivizing factors

Practically, there exists a large set of relevant incentives. Those can be often very context specific relating to the peculiarities of organizations. We aim for a more theory-driven approach in order to allow for insights that carry further than those two study contexts. More specifically, we build on the "Theory of Planned Behaviour", which is grounded in social psychology, but also well-established in other fields due to its high predictive power (e.g., Ogden, 2003; Hobbis and Sutton, 2005; McEachan et al., 2011). The framework seems particularly suitable to development economics due to its applicability to a wide variety of behaviours (e.g., Blue, 1995; Armitage and Conner, 2001; Bilic, 2005; Appleby et al., 2016) as well as within different cultural and geographical settings (e.g., Protogerou et al., 2012; Kiene et al., 2014; Hsu et al., 2017; Kassim et al., 2017).[8]

[7] We assessed in both evaluation studies whether technical knowledge or resource provision would be a main constraint, which is not the case (Diba et al., 2018; Kuhnt and Vollmer, 2018).

[8] It has to be noted that the TPB can be applied in various ways, which is likely to influence its effects (Lugoe and Rise, 1999). In order to increase the TPB's explanatory power and flexibility, several studies extend the original framework by further constructs and components (e.g., Conner and Armitage, 1998; Perugini and Bagozzi, 2001; Armitage and Conner, 2001; Cheon et al., 2012). We will stick to the original theory when applying it to development economics, while we acknowledge the propositions made to deepen or broaden the TPB. Especially, the consideration of other contextual factors offers interesting routes for further research, e.g., in the framework of the comprehensive action determination model (Klöckner and Blöbaum, 2010).

The TPB framework rests upon three determining factors that influence a person's behaviour (Fishbein and Ajzen, 1980; Ajzen, 1985). The first determinant is the personal "attitude" towards the behaviour, which refers to the degree to which a person has a favorable or unfavorable evaluation of performing the behaviour in question. A certain attitude (e.g., dis-/trust) is mostly acquired through knowledge or learning, which can be influenced by various factors, including information or previous experience (Perugini and Bagozzi, 2001; Vogel and Wanke, 2016). The second predictor termed "subjective norm" reflects the social influence felt by the individual. It refers to the perceived social pressure to perform or not to perform the behaviour. The third behavioural determinant is the degree of "perceived behavioural control," which refers to the perceived own control over the behaviour, i.e., ease or difficulty in its performance (Armitage and Conner, 2001). Generally speaking, individuals are more likely to intend a certain behaviour if they judge it beneficial (attitude toward behaviour), if they think important others want them to do it (subjective norm), and if they feel, they are able to do it (perceived behavioural control). Importantly, the TPB links its three predictors to intended behaviour, which is the immediate antecedent and, thus, a close predictor of an individual's actual behaviour (Ajzen, 1991; Bilic, 2005).

Besides the determinants suggested by the TPB, we identified one key dis-/incentivizing factor prevalent in our settings as well as in most interventions in the field of development economics: the salience of the local or international identity of the implementing agent. Higher support for a specific group of implementers could be driven by heuristics or behavioural biases, e.g., stereotypes. However, these are usually based on underlying perceptions regarding the implementers. One might not a priori prefer international or local impelementers, but support those known for higher implementation capacities, for instance. On a first view, more support for foreign implementers might be counter-intuitive as the "home bias"-phenomenon suggests

that cultural proximity could increase people's trust with regard to their assessment of the context (e.g., Fuchs and Gehring, 2017). However, an alternative strand of literature suggests increased support for foreign implementers. One reason might be the striving for high visibility among international donors (Vollmer, 2012), which is ultimately targeted to affect recipients' perceptions. Against this background, Dietrich and Winters (2015), as well as Winters et al. (2017) show that respondents link higher quality perceptions to donors rather than to the national government. Milner et al. (2016) find in Uganda that the support for foreign-funded as compared to national government funded programs is substantially larger, if participants are in favor of opposition parties, and, thus, would not be a member of the clientelist in-group. Against this background, Findley et al. (2017) stress the importance of perceptions on funding control as a main channel based on an experimental sample among Ugandan respondents.[9] 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: Firstly, an increase in contributions to impress the foreigner and, secondly, 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 shaping perceptions and attitudes and subsequent support for projects. Among the more general TPB determinants, we will consider this dis-/incentivizing aspect more closely within the broader design of our study.

Based on the TPB we formulate our first three hypotheses:

[9]Although not testing it explicitly, Findley et al. (2017) name perceptions on accountability, capacities, and level of control as further potential channels.

Hypothesis 1 (H1) *Positive attitudes towards the intervention will lead to a more supportive behaviour and, hence, increase the uptake. As attitudes are mainly shaped within an individual and less dependent on its environment, we expect them to be relevant irrespective of the context.*

Hypothesis 2 (H2) *Increased perceived behavioural control will have a positive effect on uptake. Its effect will be dependent on the level of control felt by the individual. We expect it to be more important in less controlled environments (e.g., individual health providers in Pakistan).*

Hypothesis 3 (H3) *Supportive subjective norms will lead to higher uptake. This will play a stronger role in an institutionalized, more hierarchical context, where interaction with peers is given (e.g., health facilities in Indonesia).*

With respect to the salience of local or international project implementers, we formulate our fourth hypothesis:

Hypothesis 4 (H4) *The salience of local versus international project implementers plays a role for the behaviour towards the intervention. Previous exposure to these implementing agents increases support for the intervention.*

4 Research Design and Data Collection

This paragraph describes our key measures and how we conceptualize them in our setting. Above the collection of those measures via surveys and observations, we describe the experimental design.

Measuring the concepts of the "Theory of Planned Behaviour"

In both countries we measured our data through surveys with health personnel and clinical observations of the delivery process at the end of the intervention.

The numerous applications of the TPB to a wide array of contexts ease the measurement of its determinants (e.g., French and Hankins, 2003; McEachan et al., 2011). The first determinant attitude towards the behaviour, here towards the use of the SCC, we prompt by asking the respondents to judge the usefulness of the SCC in their professional context (based upon Kam et al. (2012)). Subjective norms would translate into the degree of support by health practitioners' superiors (Sexton et al., 2006). *Perceived behavioural control* takes into account how easy the health practitioners judge the checklist to be applicable in their daily work routine. The judgment of the health practitioners on the three TPB determinants was generally very positive. For all three determinants and in both contexts the respondents provide a mean rating of five on a scale ranging from one to six, where six corresponds to "fully agree."[10] However, Appendix Tables B.1 and B.2 indicate some distinct variation, which we exploit in our analysis. Beyond the main TPB variables, surveys included demographic background information, which serves as control variables.

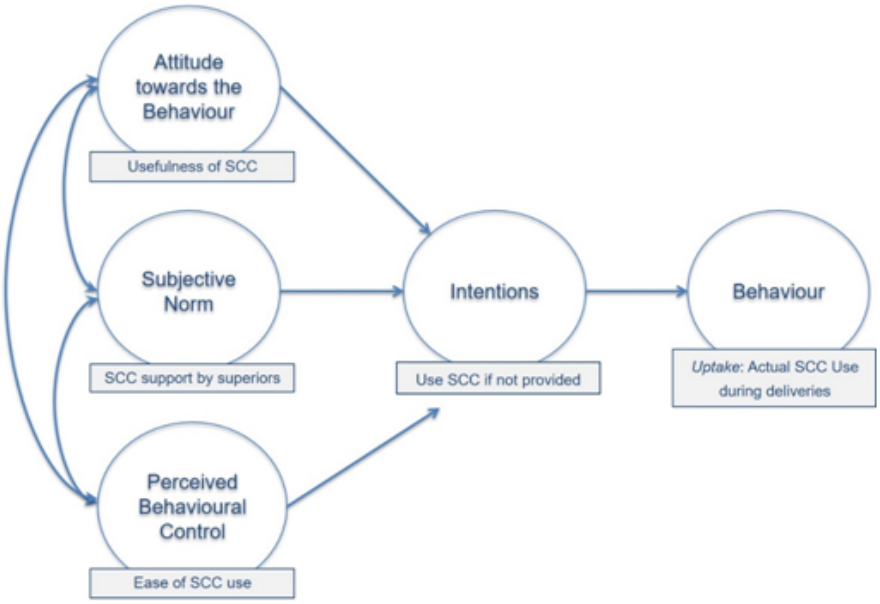
Following the TPB, the three components then influence whether health staff intends to use the checklist and, ultimately, if they actually use it during deliveries conducted (see Figure 1). Intentions to use the checklist and actual checklist use represent our outcome measures. We investigated respondents' intended behaviour towards the SCC use, by asking whether they intend to continue using the SCC after termination of the study applying a 6-point Likert scale.[11]

To also assess the actual use of the SCC, we additionally conducted standardized clinical observations in a subsample of the health facilities in Pakistan and Indonesia. Trained observers documented the delivery processes and marked whether the attending

[10] As the distribution of the TPB determinants is heavily right-skewed, we assessed robustness using a binary indicator if respondents chose the top category. Results remain robust and are available upon request.

[11] As an additional outcome measure we asked participants whether they would recommend the SCC to colleagues. Results are available upon request.

Figure 1 Applying the TPB to the SCC Intervention



Source: Authors' depiction.

Note: Own illustration based upon Ajzen (1991).

health staff had used the checklist.[12] This information was collected for 233 deliveries at 15 facilities in Indonesia and 212 deliveries at nine providers in Pakistan. We focused the observations in Pakistan on a subset of health institutions due to the organizational constraints of observing deliveries at individual practices. This difference in samples is considered when interpreting results.[13] Due to the limited number of deliveries observed per individual practitioner, we chose to aggregate the data to the provider level. Summary statistics for all measures employed can be found in Appendix

[12]Checklist use was either defined by whether the practitioners picked up the checklist during or directly after care, or whether the checklist poster was observed during the delivery process. To hang up a checklist poster in the delivery room for simultaneous consultation formed part of our intervention.

[13]In Indonesia, the fraction relates to 64 percent of all monthly conducted deliveries at observed health facilities. In Pakistan, our observations capture 50 percent of all monthly conducted deliveries at the observed health facilities as well as 94 percent of all monthly conducted deliveries at observed individual providers.

Tables B.1 for Indonesia and B.2 for Pakistan.[14]

Importantly, the data for the TPB analysis were only collected for the respondents working in treatment facilities, as at the time of the endline survey health staff in control facilities had not been in contact with the SCC. Hence, asking about the perceptions of the SCC would not have been possible and limits our sample to those interviewed at *treatment* facilities. This leaves us with 79 respondents in Pakistan and 163 health workers in Indonesia.[15] Focusing for practical reasons on the treatment facilities limits causal inference, because we face a non-random sample regarding the "dis-/incentivizing" factors suggested by the "Theory of Planned Behaviour." However, the setting of our study allows us to evaluate another dis-/incentivizing factor more closely.

4.1 The experimental set-up

As pointed out earlier, we noticed that the implementers' identity was a highly salient factor, affecting the attitude towards SCC usage. Stressing certain attributes of a particular situation among otherwise equivalent descriptions can lead to very different perceptions and behavioural reactions (Tversky and Kahneman, 1981; Kahneman, 2003; Johnson and Goldstein, 2003; Hossain and List, 2012; Payne et al., 2013). The result is what is called the *framing effect*.[16] Framed field 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 consider the following question:

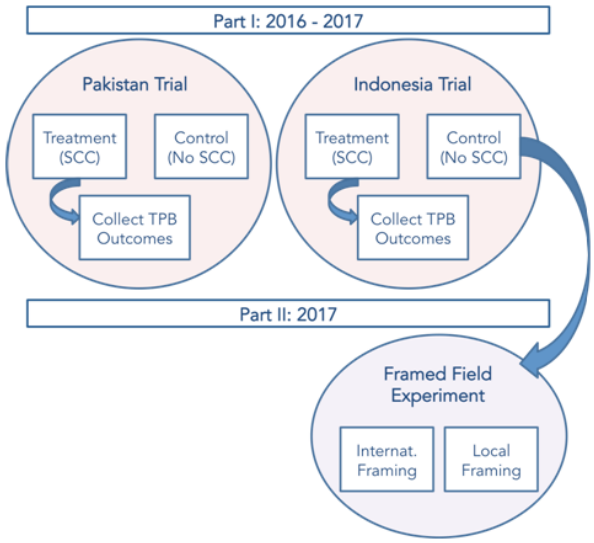
[14]More detail on the data collected can be found in Kuhnt and Vollmer (2018) and Diba et al. (2018).

[15]The Pakistani health staff worked at 70 different providers (including individual providers but also larger health facilities). While we surveyed every individual provider, we increased the number of interviews at health facilities proportionally with their number of delivery staff to get a more nuanced picture within larger teams. The Indonesian trial involved interviews at 16 health facilities.

[16]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. We consider an attribute framing, in distinction to risk or goal framings.

"Everything equal, how likely are health care 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 being aware of the de facto identity of implementers. Moreover, we would like to avoid that the framing is contaminated by heterogenous experience with checklist usage. For this reason, we make use of the treatment-control design of the SCC evaluation in Indonesia. Unfortunately, we could only consider conducting the experiment in the Indonesian sample as the tense security situation in Pakistan did not allow for additional activities.[17] The framing experiment is, thus, embedded in our study setting as described in Figure 2.

Figure 2 Study Design Flow Chart



Source: Authors' depiction.

Within the group of health practitioners working at Indonesian control facilities, we used a between-subject design and randomly assigned the study participants to two

[17]Also, due to the sampling of individual midwives in Pakistan, the organizational burden and anonymity concerns additionally prevented us from carrying out the experiment in the Pakistani context.

different framing information related to the actually conducted SCC intervention.[18] The first framing information stressed the involvement of international actors in the intervention, while the second made the participation of local counterparts more salient.[19]

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.[20] 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$). Afterwards, the enumerators offered the respondents to participate in the experiment.[21] The "experimental commodity" was derived from the on-going larger SCC intervention. First, the idea and structure of the SCC was explained to the participants. Afterwards, 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.

We then investigated the participants' respective behaviour towards 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

[18]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.

[19]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.

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

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

support the implementation of the SCC in other anonymous health facilities within the province.[22] The monetary contribution was directly deducted from the voucher for phone credit top-up in appreciation of their survey participation.[23] 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 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 towards 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.[24]

In order to get a clearer understanding how previous experience 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

[22]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, while the additional outcomes are presented in Section Appendix.1.

[23]If respondents wanted to contribute, we offered them five options from 5,000 to 25,000 IDR (equivalent to 0.4 – 1.9 US$) due to pragmatic reasons of specific top-up values.

[24]For the correlations, please see Table B.15. We did not use those channels for further analysis. Yet, they might be still informative in terms of general attribute ascription.

on experience with local and international agents.

For the detailed experimental protocol including the specific framing information see Appendix.2.

5 Empirical approach and descriptive data

Empirical approach

In the first part of our regression analysis we address the role of the "dis-/incentivizing" factors for intended behaviour with regard to checklist use. Our regression line for intended behaviour reads as follows:

$$y_i = \alpha + \beta_i TPBdeterminant_i + \beta_k \sum_k X_i + \epsilon_i \quad (1)$$

As throughout the study, we estimate models for Indonesia and Pakistan separately using ordinary least squares (OLS) regressions. Our level of analysis is the individual health worker i (79 respondents for Pakistan and 163 individuals for Indonesia). y_i determines our outcome variable, which measures intended behaviour employing 6-point Likert scales. α is a constant, and $TPBdeterminant_i$ captures our variables of interest (also using 6-point Likert scales) via our three perception measures for the three TPB pillars: *Attitudes, subjective norms,* and *perceived behavioural control*.

We employed Likert scales to all perception-based survey questions, which are relatively continuous measures. Hence, we consider them as continuous variables in the estimations, which is the preferred method of analysis proposed in the literature (Pasta, 2009).[25] As our sample is restricted to our treatment group and includes, thus,

[25]Further, we also estimated regressions with an alternative coding for robustness, where we defined a dummy variable with the value one for the highest category and zero otherwise. Results are robust and available upon request. In a pre-trest we also assessed the feasibility of continuous items with a scale from 0 to 100, but learned that those were harder to comprehend for respondents.

a non-random set of individuals, estimations are not derived within the randomization framework and do not allow a causal interpretation. Nonetheless, controlling for several potentially confounding variables, we will receive informative correlations about how behavioural processes are associated with intervention uptake. In adjusted regressions we add $\sum_k X_i$, which represents our set of k control variables. These include a binary variable indicating the location of the facility (rural versus urban), a variable capturing the district where the provider is located, the level of service provision, which is proxied by a dummy for 24/7 opening hours, and a variable indicating the type of facility.[26] Those time-invariant facility characteristics might be both correlated with the drivers of the TPB as well as the outcome and, hence, cause omitted variable bias if not considered. Perceived behavioural control could be affected by staffing and equipment, which is captured by the facility type and geographical remoteness (district dummies and rural/urban distinction) as well as the 24/7 service provision. Remoteness, services and facility type also influence the safety culture, which affects providers' attitudes and the subjective norms of superiors towards the SCC.

Our second part of regressions is the equivalent to the first but changes the outcome variable to birth observations i measuring the actual behaviour. Here, y_i, is a binary variable equalling one, if the checklist was used by the health worker during the delivery. As we cannot link each delivery to the specific health workers' responses, we take averages of *attitudes, subjective norms* and *perceived behavioural control* per health facility. Those averages provide us with an intuition of more supportive environments

[26]This variable captures the different types of providers (which are more general than facility dummies). Our sample includes a wide heterogeneity of facilities from primary to tertiary health providers where this variable captures their specificities, including team size, resource access, or delivery load. Research from different facility types indicates very heterogeneous uptake and different *attitudes* of the respondents towards the tool (Semrau et al., 2017; Kabongo et al., 2017; World Health Organization, 2018).

being associated with more or less take-up.[27] The control variables X_j stay the same as in regression line (1).

The third part of our regression analysis concerns the experimental data. Our analysis of the framed field 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:

$$y_i = \alpha + \beta_1 framing_i + \beta_2 framing_i * c_i + \beta_3 c_i + \beta_m \sum_m C_i + v_i \tag{2}$$

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. Moreover, heterogeneous effects are assessed by the inclusion of an interaction between the framing and c_i, which is prior participation in international or local projects. We are, thus, mainly interested in the effect sizes of β_1 and β_2.

While the randomization ensures exogeneity of the framing, project participation is potentially endogenous regarding other traits of the surveyed respondent. However, as recent 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.[28]

[27] As our analysis, thus, involves different aggregation levels and our measures of intention and actual behaviour capture slightly different concepts, we do not estimate a model on the direct link between intentions and behaviour.

[28] 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 B.6 and B.7 underscore that previous participation is balanced across both framing treatments.

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 binary variable marking whether the respondent experienced financial problems within the past days as this might affect monetary contributions.[29] Further, to control for a potential social desirability bias, we measured social conformity following the social desirability scale developed by Kemper et al. (2014). This measure was adopted to the Acehnese context and we transformed its five items into a composite index.[30] We control also for the subjective perception regarding the amount of paperwork during deliveries, which was motivated by an often-experienced perception during implementation that the new tool adds to the already existing paperwork. Finally, v_i describes the residual. Errors are clustered at the facility level to take into account similarities within teams.[31]

Descriptive data

Descriptive statistics show that the SCC is generally valued by the practitioners in Indonesia and Pakistan (Figure B.3). Yet, there is some distinct variation within and across the settings. Additionally, Figure B.4 describes the actual SCC use by health practitioners in Indonesia and Pakistan. It indicates a limited uptake and, hence, a potential gap between intended and actual use.

[29]Related research has similarly controlled for a constructed wealth index (e.g., Cilliers et al., 2015).

[30]We adapted the social desirability measures to the respective context in cooperation with Indonesian counterparts. For instance, one of the items reads "I have occasionally thrown litter away in the countryside or on to the road." As environmental concerns are less salient in the Acehnese context than religious concerns, we changed the item to "When I had the chance to donate for religious purposes, I always contributed a lot." The full set of questions we used for the construction of the social desirability index are displayed in Appendix.2.

[31]Due to a limited number of clusters we also present results with wild bootstrapped standard errors following Cameron et al. (2008) for all our baseline models in the Appendix. However, this is only possible for the unadjusted regressions (without controls). When bootstrapping standard errors in models with control variables, we face problems of overfitting. This is the case as our controls consist mainly of dummy or categorical variables, which reduce variation among our relatively small number of observations too strongly to calculate meaningfully adjusted standard errors. Accordingly, we prefer to present regressions without bootstrapped standard errors in our main models.

Experimental data Individual characteristics and further contextual variables are balanced across framings indicating that the randomization was successful (Appendix Table B.6). In our main analysis, we focus on those participants that have not been in prior contact with the SCC as 27.92 percent of the respondents state that they were previously exposed to the SCC.[32] 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.[33] This reduces our sample to 173 participants.[34] Balance on important covariates is also given in this reduced sample (see Appendix Table B.7). 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 percent (17.5 percent) 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 percent of the respondents faced international research projects in their facility during the previous two years.

[32]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 percent 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.

[33]As a robustness check, we also report the full sample results including a prior contact binary variable in the regression model in Appendix Table B.12. However, as we assume a large heterogeneity of exposure – health practitioners with a job at another facility might have worked with the SCC, others might have just heard the name of the SCC from colleagues – we prefer the reduced sample for our main results.

[34]Due to two outcome measures that could not be matched to respondents and four respondents that refrained from answering on control questions, the sample is reduced to n=165 in our main specifications.

6 Results

Main results: TPB determinants and SCC uptake

For all three TPB determinants, *attitudes*, *subjective norms*, and *perceived behavioural control*, in both study sites, we find that coefficients consistently point towards a positive direction.[35] Tables 1 and 2 display the regression results of the intended and actual SCC uptake for the data from Pakistan and Indonesia. While the first row always presents the unadjusted coefficients, the second displays results adjusted for control variables as described in Section 5. Results show that respondents who express a strongly positive *attitude* towards the SCC are also more likely to intend to use the new tool even if it is not freely provided to them anymore (columns (1a) to (2b)). In Pakistan and Indonesia the coefficients are positive and statistically significant (ranging from the 1 percent to 5 percent level).

[35]In order to get a notion of the explanatory power of each TPB determinant, we introduce the concepts separately. Taking into account the interconnectedness of the three variables, we considered an index based on principal component analysis as a robustness test. Results in Table B.5 suggest a robust positive relation of the index both with intentions and behaviour.

Table 1 Theory of Planned Behaviour – Intended SCC uptake

	Intended Behaviour			
	Would use SCC even if copies are not provided			
	1 "disagree strongly" – 6 "agree strongly"			
	Pakistan		Indonesia	
	(1a)	(1b)	(2a)	(2b)
Attitudes:				
SCC in professional role: 1 "completely useless" – 6 "completely useful"				
	0.984***	0.818***	0.454***	0.309**
p-value	(0.000)	(0.000)	(0.004)	(0.012)
Adjusted R^2	0.187	0.254	0.114	0.272
N	79	79	163	163
Subjective Norms:				
SCC is supported by superiors: 1 "not at all" – 6 "completely"				
	0.143	0.164*	0.536***	0.316***
p-value	(0.115)	(0.060)	(0.007)	(0.001)
Adjusted R^2	0.008	0.304	0.132	0.261
N	58	58	163	163
Perceived Behavioural Control:				
Ease of SCC in work environment: 1 "very difficult" – 6 "very easy"				
	0.439***	0.366**	0.261*	0.023
p-value	(0.003)	(0.029)	(0.090)	(0.863)
Adjusted R^2	0.128	0.211	0.048	0.222
N	78	78	163	163
Control variables	No	Yes	No	Yes
Mean of dep. var.	4.628	4.628	4.847	4.847
Median of dep. var.	5	5	5	5
SD of dep. var.	1.452	1.452	0.634	0.634

Note: All regressions are based upon the treated providers. Adjusted regressions (b) additionally control for a variable indicating the facility type, a binary variable indicating rural/urban location, a variable indicating the district and a binary variable indicating whether the facility is open 24/7. Standard errors (SE) are clustered at the facility level. Asterisks indicate p-values according to: * p<0.1, **p<0.05, *** p<0.01.

Table 2 Theory of Planned Behaviour – Actual SCC uptake

	Actual Behaviour Was SCC actively used or looked at during delivery? 0 "No" – 1 "Yes"			
	Pakistan		Indonesia	
	(1a)	(1b)	(2a)	(2b)
Attitudes:				
SCC in professional role: 1 "completely useless" – 6 "completely useful"				
	0.655***	0.471**	-0.356	0.394***
p-value	(0.003)	(0.020)	(0.245)	(0.000)
N	212	212	219	219
Subjective Norms:				
SCC is supported by superiors: 1 "not at all" – 6 "completely"				
	0.207*	0.078**	0.654*	0.279***
p-value	(0.097)	(0.027)	(0.091)	(0.000)
N	212	212	219	219
Perceived Behavioural Control:				
Ease of SCC in work environment: 1 "very difficult" – 6 "very easy"				
	0.306***	0.112	0.059	0.015
p-value	(0.000)	(0.169)	(0.423)	(0.979)
N	212	212	219	219
Control variables	No	Yes	No	Yes
Mean of dep. var.	0.344	0.344	0.389	0.389
SD of dep. var.	0.476	0.476	0.489	0.489

Note: All regressions are based upon the treated providers. Adjusted regressions (b) additionally control for a variable indicating the facility type, a binary variable indicating rural/urban location, a variable indicating the district and a binary variable indicating whether the facility is open 24/7. Standard errors (SE) are clustered at the facility level. Asterisks indicate p-values according to: * p<0.1, **p<0.05, *** p<0.01.

This is also supported by the actual SCC use (in Table 2 columns (1a) to (2b)). The stronger the positive stance towards the checklist, the more often health staff actively uses the SCC during the delivery process. If the SCC is perceived to be more useful (*attitude*), its actual use among Indonesian health workers increases by 39.4 percentage points and among Pakistani practitioners by 47.1 percentage points.

Further, we find consistently positive coefficients in both countries with respect to the support of superiors for the new tool (*social norms*). While it seems to play an important role for intended and actual SCC uptake in Indonesia, it is less important for intended behaviour as compared to the actual SCC use in the Pakistani setting. Considering the different samples across intentional and behavioural outcome measures in Pakistan help to interpret those results.

In both countries, Pakistan and Indonesia, the actual behavioural outcome was mainly collected for health practitioners working in facilities. In these facilities hierarchical structures are dominant and the stance of the superiors towards the SCC is more critical. While we measured intentions to use the SCC also mainly in facilities in Indonesia, the majority of respondents in Pakistan on intentions are individual health workers (such as community midwives). They work alone without direct supervision and are not integrated into a hierarchically structured team. Hence, for them the opinion of superiors is less of a concern but rather the perceived usability (*perceived behavioural control*). In this regard, we see that the ease of use is a statistically significant predictor of intended SCC use in Pakistan (at the 5 percent level in the adjusted regression), while it is positive but not statistically significant in the Indonesian context or for actual SCC uptake in both countries.[36] These results – though not allowing the establishment of a causal pathway – give a consistent indication: Influencing the TPB determinants into the respective positive direction, is associated with increased intended and actual uptake of the SCC.

Differences in the significance across TPB determinants are well in line with qualitative evidence. Indonesian coaches, who assisted health personnel in using

[36]As outlined above, we use wild cluster bootstrapped standard errors as robustness tests in samples with a small number of clusters (9 in Pakistan and 15 in Indonesia). Results are displayed in Appendix Table B.4 showing that results are by and large robust to this standard error adjustment. When we generate a dummy variable as an outcome, equaling one for the highest category only (thus, if respondents "fully agree" to "Would try to use SCC even if copies are not provided") results are qualitatively unchanged (see Appendix Table B.3).

the checklist, were seldomly asked for help regarding the content of the SCC, which corresponds to the ease of use of this intervention. In contrast, the assessment of the supervisor seems to matter a lot in the Indonesian society, where workplace position and seniority play a predominant role. This is also borne out by inter-facility staff meetings and midwives' correspondence with coaches in Indonesia, stressing the salience of supervisors and colleagues reminding each other to use the checklist regularly. In the Pakistani case, we see a stronger relationship with *attitudes* and *control* rather than with *norms*. In line with explanations from above, the effect is likely to be driven by the sample of community midwives, who work rather self-employed and do not depend on superiors' norms, accordingly.[37]

Both sets of results imply that in both countries, specifically, *attitudes* are crucial in shaping intentions and actual behaviour. This is in line with our Hypothesis 1. While *social norms* as well as *control* are both positiviely related to uptake in both countries, we find that as expected in Hypotheses 2 and 3 both determinants are more context dependent. Analyzing the data from the framed experiment, in the subsequent section we investigate whether the implementer's background acts as another important dis-/incentivizing factor for determining support towards an intervention.

Main results: framing experiment

Table 3 displays the main results of the framing experiment conducted in Indonesia. We only include our main outcome measure (monetary investment) here, while results of the alternative outcomes are presented in the Appendix (Table B.10). The first column presents the unadjusted results, whereas the second column gives the results

[37]Community midwives in Pakistan are trained midwives, employed by the district governments, who operate on their own within local, often rural, communities. They are only loosely attached (e.g., through provision of medication and equipment) to the local government structures and de facto there is very little oversight of their activities.

adjusted for additional control variables.[38] We limit our sample to those respondents who were not exposed to the SCC prior to this experiment (see 5). Full sample regression results controlling for prior contact, are shown in the Appendix (Table B.12) and are comparable to the findings presented in the main part.[39] As a conservative robustness check, we also present random inference based p-values.[40] 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 5 percent 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 1,284 IDR higher.[41]

[38] In line with the randomized setup of the study, results are robust to the inclusion of further covariates, which increases the precision of estimates. The full specification including all control variables is presented in the Appendix Table B.9.

[39] As a further robustness check we estimate a regression, which controls for an interaction of the framing with the indicator for past contact. 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 to the SCC and induce, thus, a downward bias. Second, individuals with prior contact to the checklist might react heterogeneously due to more comprehensive information. Table B.13 depicts the corresponding results. While the framing indicator decreases slightly in size, but stays significant in the adjusted regressions, there is no significantly different treatment effect for those respondents with past contact.

[40] 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.

[41] These results are supported by the alternative outcome measures presented in Appendix Table B.10.

Table 3 Framing Experiment – Main Results

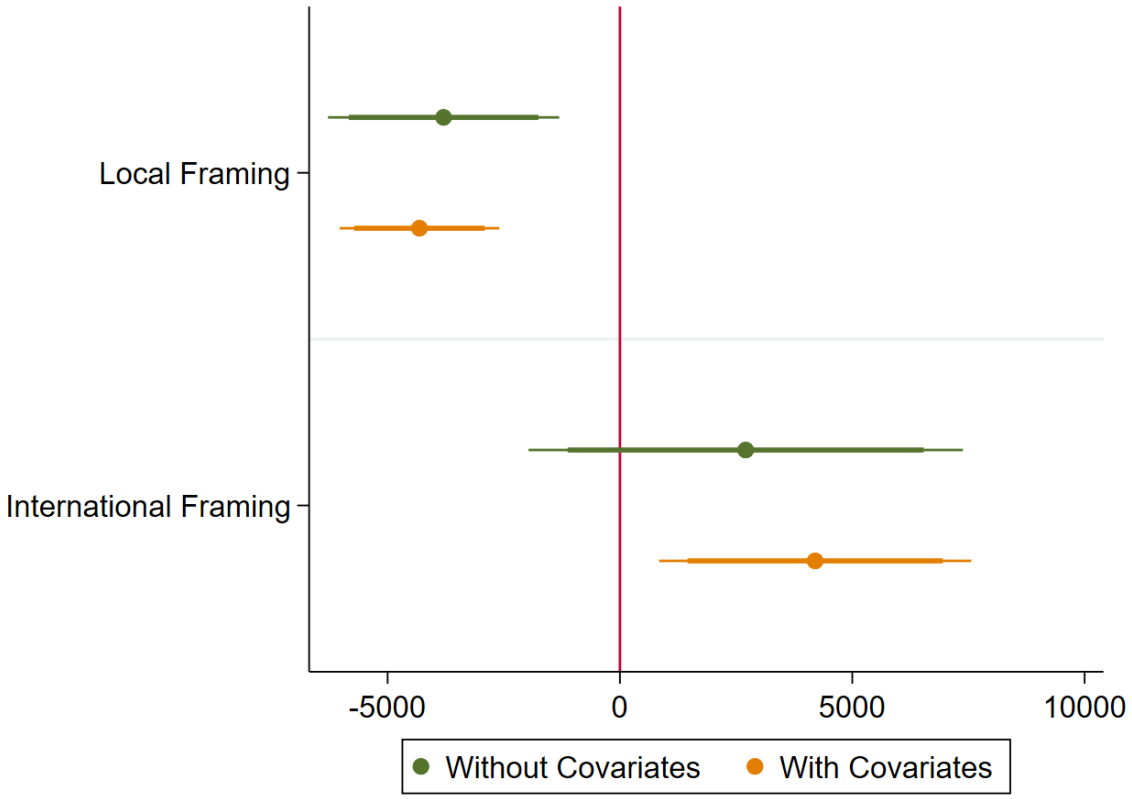
Financial Contribution in support of SCC project (in IDR)		
	(a)	(b)
Framing: 1 = "international"	557.6236	1,283.7717**
p-value	(0.396)	(0.021)
RI p-value	(0.450)	(0.057)
N	165	165
Control variables	No	Yes
Mean of dep. var.	4,757.576	4,757.576
SD of dep. var.	4,711.366	4,711.366

Note: All specifications are based upon the sample limited to those respondents without prior SCC contact. Specifications (b) 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. The same regression with wild cluster bootstrapped SE can be found in Appendix Table B.8, 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<0.1, **p<0.05, *** p<0.01.

In order to understand in more detail why respondents show stronger support towards projects implemented by international actors as compared to local implementers, we investigated *previous exposure* as a mechanism that is likely to influence the behaviour of respondents. Previous exposure is one prominent factor shaping ideas and attitudes. Hence, it might play a role 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 respective actors. Aceh is specific due to the activity of manifold – oftentimes international – donors in response to the human tragedy of 2004's tsunami.

Figure 3 displays the point estimates and confidence intervals for the interaction

Figure 3 Framing Experiment – Previous Experience



Note: This Figure depicts coefficients for tripple interactions of the local/international framing with indicators of participation in international projects. Covariates 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. 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 prior experience with either actor and faced a local framing. The corresponding point estimates are depicted in Table B.16. 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.

of our experimental framing with the binary variables indicating if respondents have already participated in international or local research projects. In order to facilitate interpretation 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. Interpretation, thus, focuses on this group, while complete results are presented in Appendix Table B.16. As before, the framing indicator equals one for the international framing treatment and zero for the local framing treatment.

Green bars in Figure 3 indicate the coefficients of regressions without covariates and orange bars the adjusted point estimates. Regarding confidence intervals, thick bars refer to the 10 percent and thin bars to the 5 percent 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 6,500-8,500 IDR (e.g., 0.45-0.65 US$) if those health workers face the local framing (p-value: 0.023 without control variables; p-value: 0.000 with control variables).[42] In contrast, this implies that the attitude towards the intervention is significantly more positive if respondents knowing both international and local actors are framed internationally. For respondents with international and local experience we find the only significant group-wise difference between individuals with comparable experience.

Thus, the results from Figure 3 suggest in line with our Hypothesis 4 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.

[42] Although this amount seems small, it corresponds to one meal or half an hour of work of a health worker in the local context.

Qualitative research Qualitative data based on 66 surveys with health practitioners was collected to provide a clearer understanding 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 towards internationals are mostly linked to experiences of more structured implementation approaches (13 indications) and a higher perceived level of knowledge (4 indications). Moreover, 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 and 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 B.15) [43]

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

[43]We asked health practitioners if they would attribute certain characteristics rather to local or international researchers (e.g., skills, corruption, 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, indicated by the significant framing effects in Table B.15. Hence, we did not use those channels for further analysis. Yet, they might be still informative in terms of general attribute ascription.

7 Discussion and Conclusion

Many interventions in the field of development economics suffer from low uptake by the target population. Supportive behaviour, however, is a crucial ingredient for a successful intervention. But what factors influence/explain heterogeneous uptake by the target population? Evidence from behavioural economics supports the importance of non-monetary incentives, trust, or peer effects to explain human behaviour. These insights are also of utmost importance to the design of interventions in development economics as the majority of these projects aims at changing human behaviour. Our study focuses on investigating dis-/incentivizing factors explaining variations in uptake by the target population. Three important mechanisms to explain human behaviour (in reaction to an intervention) are proposed by the *Theory of Planned Behaviour (TPB)* – a well-established theory originating from social psychology. The TPB offers a systematic approach to explain and influence supportive human behaviour by considering three determinants: A positive *attitude* towards the behaviour or intervention, supporting *subjective norms*, and a high degree of *perceived behavioural control*.

We study these determinants within the framework of an intervention implemented in Pakistan and Indonesia. Using the introduction of a new tool, the Safe Childbirth Checklist (SCC) among health practitioners, we provide evidence of the positive association of all three mechanisms with the intended and actual uptake of the SCC in both country settings. A more positive *attitude* towards the new tool (the SCC) is associated significantly with increased intended and actual use of the SCC in both cultural contexts. While *subjective norms* in favor of the intervention are particularly important in larger health facilities in Indonesia characterized by more pronounced hierarchies, greater *perceived behavioural control* to actively use and implement the checklist is a more important determinant among individually working

health practitioners in Pakistan. It is important to note that this analysis does not allow us to infer causal effects, although we condition our analysis on a broad set of confounding factors. Studying dis-/incentivizing factors of a similar intervention in two diverse study contexts strengthens the claim of generalizability of the results. Previous studies on the determinants of the TPB also support its broad applicability to explain and influence human behaviour.

Following recent evidence we were able to study another potentially critical dis-/incentivizing factor for human behaviour towards interventions in Indonesia: The implementer's background. More specifically, we investigate how the salience of a local versus an international agent causally influences the participants' support for the project. This is of particular interest as the majority of interventions in the field of development economics are cooperations between local and international agents. The results of the framed field experiment indicate that respondents are more supportive towards interventions (measured through monetary support) implemented by international actors as compared to solely locally led projects. This finding is in line with previous research on behavioural reactions towards international and multilateral donor agencies (e.g., Milner et al., 2016; Winters et al., 2017).

Our results suggest that previous experience is pivotal. Those respondents that have already been exposed to previous internationally-led research interventions take a more positive stance towards 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 (Doocy et al., 2007) facilitates the assessment towards the different implementers. However, this context of ultimate human emergency, might have induced a more positive

stance towards the international assistance and makes the interpretation specific to the context.[44]

Many high-level fora voiced demands for a higher effectiveness of global development cooperation, including the Paris Declaration (2005), the Accra Agenda for Action (2008) and the Busan Partnership for Effective Development Co-Operation (2011). Local uptake is a crucial prerequisite for more effective actions towards global sustainable development. Our study provides evidence for the importance of considering dis-/incentivizing factors when aiming at influencing the uptake of interventions. Our results suggest an active consideration of the TPB determinants in the design and implementation of interventions to positively affect uptake by the targeted population. Certainly, researchers and practitioners will already have intuitively taken determinants of the TPB into account when designing their intervention. In our study, however, we argue for a systematic application of the TPB to increase uptake rates, an important ingredient for the success of a project. A qualitative investigation prior to the project implementation and close cooperation with people knowing the local context to identify behavioural, normative, and control beliefs (that underlie the TPB determinants) within the study sample is recommended (Protogerou et al., 2012; Hobbis and Sutton, 2005). Our results regarding the salience of international versus local project implementers have to be considered against the background of the respective local context, which defines the previous exposure to implementing agents. Generally, using framing as a tool to make a well-regarded implementing agent more salient might be a low-hanging fruit to increase supportive behaviour of population groups in a cost-effective way (Bertrand et al., 2006). In order to increase local ownership it is of utmost importance to generate positive experiences of the target group with locally

[44]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. Moreover, Aceh might be specific due to its strong Muslim heritage and introduction of Islamic law in 2006.

led projects. Here, strengthening capacities of local agencies is necessary, also to foster later scalability of projects led by the local government. 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.

While this study uses the setting of a research project in the field of maternal and child health and this involving specific intervention characteristics, we believe that our results are likely to be also valid for programs implemented by practical development cooperation and within the context of other areas of development interventions. Further research needs to contribute to a clearer understanding of the potential effects of dis-/incentivizing factors and more specifically the TPB on human behaviour by randomly altering these determinants or replicating results in different settings. This way, important knowledge can be gained to not only increase uptake of research interventions, but also of practical development cooperation.

References

Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. In J. Kuhl and J. Beckmann (Eds.), *Action-control: From cognition to behavior*, Chapter 2, pp. 11–39. Heidelberg, Berlin: SSSP Springer Series in Social Psychology.

Ajzen, I. (1991). The theory of planned behavior. *Organisational Behaviour and Human Decision Processes 50*, 179–211.

Appleby, B., C. Roskell, and W. Daly (2016). What are health professionals' intentions toward using research and products of research in clinical practice? A systematic review and narrative synthesis. *Nursing open 3*(3), 125–139.

Armitage, C. J. and M. Conner (2001). Efficacy of the Theory of Planned Behaviour: A meta-analytic review. *British Journal of Social Psychology 40*, 471–499.

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

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

BBC (2005). Germany takes lead in aid effort. Available at http://news.bbc.co.uk/2/hi/europe/4148719.stm, last accessed October 4, 2018.

Behavioral Science Solution (2018). Framing Effect. Available at https://www.behavioraleconomics.com/resources/mini-encyclopedia-of-be/framing-effect, last accessed September 29, 2018.

Berthélemy, J.-C. (2006). Bilateral donors' interest vs. recipients' development motives in aid allocation: Do all donors behave the same? *Review of Development Economics 10*(2), 179–194.

Bertrand, M., S. Mullainathan, and E. Shafir (2006). Behavioral Economics and Marketing in Aid of Decision Making Among the Poor. *Journal of Public Policy & Marketing 25*(1), 8–23.

Bilic, B. (2005). The Theory of Planned Behaviour and Health Behaviours: Critical Analysis of Methodological and Theoretical Issues. *Hellenic Journal of Psychology 2*, 243–259.

Blue, C. (1995). The Predictive Capacity of the Theory of Reasoned Action and the Theory of Planned Behavior in Exercise Research: An integrated Literature Review. *Research in Nursing and Health 18*, 105–121.

Borchard, A., D. L. B. Schwappach, A. Barbir, and P. Bezzola (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. and S. Polania-Reyes (2012). Economic incentives and social preferences: substitutes or complements? *Journal of Economic Literature 50*(2), 368–425.

Bun, M. J. and T. D. Harrison (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. and D. Dollar (2000). Aid, Policies, and Growth. *American Economic Review 90*(4), 847–868.

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

Cameron, D. B., A. Mishra, and A. N. Brown (2016). The growth of impact evaluation for international development: How much have we learned? *Journal of Development Effectiveness 8*(1), 1–21.

Cheon, J., S. Lee, S. M. Crooks, and J. Song (2012). An investigation of mobile learning readiness in higher education based on the theory of planned behavior. *Computers and Education 59*(3), 1054–1064.

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

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

Conner, M. and C. J. Armitage (1998). Extending the Theory of Planned Behavior: A Review and Avenues for Further Research. *Journal of Applied Social Psychology 28*(15), 1429–1464.

Croskerry, P. (2002). Achieving quality in clinical decision making: cognitive strategies and detection of bias. *Academic Emergency Medicine 9*(11), 1184–1204.

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

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

Diba, F., I. Ichsan, M. Muhsin, M. Marthoenis, H. Sofyan, S. Susanti, M. Andalas, K. Richert, L. Kaplan, Samadi, and S. Vollmer (2018). Analysis of the Impact and Effectiveness of Safe Childbirth Checklists on the Quality of Care and Birth Outcomes in Health Facilities in Aceh Province, Indonesia – A Randomized Controlled Trial. Mimeo.

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

Dollar, D. and L. Pritchett (1998). Assessing Aid – What Works, What Doesn't, and Why. World Bank Policy Research Report 1.

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

Duflo, E., R. Glennerster, and M. Kremer (2007). Chapter 61 Using Randomization in Development Economics Research: A Toolkit. *Handbook of Development Economics 4*(07), 3895–3962.

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

Findley, M. G., A. S. Harris, H. V. Milner, and D. L. Nielson (2017). Who Controls Foreign Aid? Elite versus Public Perceptions of Donor Influence in Aid-Dependent Uganda. *International Organization* (October 2015), 1–31.

Fishbein, M. and I. Ajzen (1980). *Understanding Attitudes and Predicting Social Behavior* (Revised ed.). Prentice-Hall.

French, D. P. and M. Hankins (2003). The expectancy-value muddle in the theory of planned behaviour – and some proposed solutions. *British Journal of Health Psychology 8*, 37–55.

Fuchs, A. and K. Gehring (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., S. Meier, and P. Rey-Biel (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives 25*(4), 191–210.

Haugen, A. S., E. Søfteland, S. K. Almeland, N. Sevdalis, B. Vonen, G. E. Eide, M. W. Nortvedt, and S. Harthug (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.

Hobbis, I. C. A. and S. Sutton (2005). Are Techniques Used in Cognitive Behaviour Therapy Applicable to Behaviour Change Interventions Based on the Theory of Planned Behaviour? *Journal of Health Psychology 10*(1), 7–18.

Hoffman, M., V. Mukete, J. Szmanda, and C. Witt (2011). The effect of increased stress levels on memory test performance. Unpublished manuscript available at https://www.google.com/search?q=The+effect+of+increased+stress+levels+on+memory+test+performance+hoffmann&ie=utf-8&oe=utf-8&client=firefox-b-1-ab.

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

Hsu, C. L., C. Y. Chang, and C. Yansritakul (2017). Exploring purchase intention of green skincare products using the Theory of Planned Behavior: Testing the moderating effects of country of origin and price sensitivity. *Journal of Retailing and Consumer Services 34*, 145–152.

Johnson, E. J. and D. Goldstein (2003). Do Defaults Save Lives? *Science 302*(5649), 1338–1339.

Kabongo, L., J. Gass, B. Kivondo, N. Kara, K. Semrau, and L. R. Hirschhorn (2017). Implementing the WHO Safe Childbirth Checklist: lessons learnt on a quality improvement initiative to improve mother and newborn care at Gobabis District Hospital, Namibia. *BMJ Open Quality 6*(2), 1–8.

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *The American Economic Review 93*(5), 1449–1475.

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica 47*(2), 263–291.

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

Kam, L., V. E. Knott, C. Wilson, and S. K. Chambers (2012). Using the Theory of Planned Behavior to understand health professionals' attitudes and intentions to refer cancer patients for psychosocial support. *Psycho-Oncology 21*, 316–323.

Karim, A. M., T. Williams, L. Patykewich, D. Ali, C. E. Colvin, J. Posner, and G. Rutaremwa (2009). The impact of the African youth alliance program on the sexual behavior of young people in Uganda. *Studies in Family Planning 40*(4), 289–306.

Kassim, K. A., L. Arokiasamy, M. H. M. Isa, and C. H. Ping (2017). Intention to Purchase Safer Car: an Application of Theory of Planned Behavior. *Global Business & Management Research 9*(1), 188–197.

Kemper, C. J., C. Beierlein, D. Bensch, A. Kovaleva, and B. Rammstedt (2014). *Soziale Erwünschtheit-Gamma (KSE-G) (Version 1.0). ZIS - GESIS Leibniz Institute for the Social Sciences.* GESIS.

Kettle, S., M. Hernandez, S. Ruda, and M. Sanders (2016). Behavioral interventions in tax compliance: evidence from Guatemala. World Bank Policy Research Working Paper 7690.

Kiene, S. M., S. Hopwood, H. Lule, and R. K. Wanyenze (2014). An empirical test of the Theory of Planned Behaviour applied to contraceptive use in rural Uganda. *Journal of Health Psychology 19*(12), 1564–1575.

Kinsman, J., J. Nakiyingi, A. Kamali, L. Carpenter, M. Quigley, R. Pool, and J. Whitworth (2001). Evaluation of a comprehensive school-based AIDS education programme in rural Masaka, Uganda. *Health Education Research 16*(1), 85–100.

Klöckner, C. A. and A. Blöbaum (2010). A comprehensive action determination model:

Toward a broader understanding of ecological behaviour using the example of travel mode choice. *Journal of Environmental Psychology 30*(4), 574–586.

Kuhnt, J. and S. Vollmer (2018). Implementing the WHO safe childbirth checklist in Pakistan: Can a simple checklist reduce maternal and neonatal health outcomes? Mimeo.

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

Lugoe, W. and J. Rise (1999). Predicting Intended Condom Use among Tanzanian Students using the Theory of Planned Behaviour. *Journal of Health Psychology 4*(4), 497–506.

Madrian, B. C. (2014). Applying Insights from Behavioral Economics to Policy Design. NBER Working Paper 20318.

McEachan, R. R. C., M. Conner, N. J. Taylor, and R. J. Lawton (2011). Prospective prediction of health-related behaviours with the Theory of Planned Behaviour: A meta-analysis. *Health Psychology Review 5*(2), 97–144.

Milner, H. V., D. L. Nielson, and M. G. Findley (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.

Minasyan, A., P. Nunnenkamp, and K. Richert (2017). Does Aid Effectiveness Depend on the Quality of Donors? *World Development 100*, 16–30.

Nayiga, S., D. DiLiberto, L. Taaka, C. Nabirye, A. Haaland, S. G. Staedke, and C. I. Chandler (2014). Strengthening patient-centred communication in rural

Ugandan health centres: A theory-driven evaluation within a cluster randomized trial. *Evaluation 20*(4), 471–491.

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

Ogden, J. (2003). Some problems with social cognition models: a pragmatic and conceptual analysis. *Health Psychology 22*, 424–428.

Pasta, D. J. (2009). Learning when to be discrete: Continuous vs. categorical predictors. In *SAS Global Forum*, Number 248.

Payne, J. W., N. Sagara, S. B. Shu, K. C. Appelt, and E. J. Johnson (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.

Perugini, M. and R. P. Bagozzi (2001). The role of desires and anticipated emotions in goal-directed behaviours: Broadening and deepening the theory of planned behaviour. *40*, 79–98.

Protogerou, C., A. J. Flisher, L. E. Aar, and C. Mathews (2012). The theory of planned behaviour as a framework for predicting sexual risk behaviour in sub-Saharan African youth: A critical review. *Journal of Child and Adolescent Mental Health 24*(1), 15–35.

Rajan, R. G. and A. Subramanian (2008). Aid and Growth: What Does the Cross-Country Evidence Really Show? *Review of Economics and Statistics 90*(4), 643–666.

Rogers, P. (2014). Theory of change. UNICEF Methodological Briefs 2, UNICEF. Available at https://www.unicef-irc.org/publications/

747-theory-of-change-methodological-briefs-impact-evaluation-no-2.html, last accessed September 27, 2018.

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

Semrau, K. E., L. R. Hirschhorn, M. Marx Delaney, V. P. Singh, R. Saurastri, N. Sharma, D. E. Tuller, R. Firestone, S. Lipsitz, N. Dhingra-Kumar, et al. (2017). Outcomes of a coaching-based WHO Safe Childbirth Checklist program in India. *New England Journal of Medicine 377*(24), 2313–2324.

Sexton, J. B., R. L. Helmreich, T. B. Neilands, K. Rowan, K. Vella, J. Boyden, P. R. Roberts, and E. J. Thomas (2006). The safety attitudes questionnaire: psychometric properties, benchmarking data, and emerging research. *BMC health services research 6*(1), 44.

Simon, H. A. and J. Feldman (1959). Theories of Decision-Making in Economics and Behavioral Science. *Source: The American Economic Review 49*(3), 253–283.

Toll, B. A., S. S. O'Malley, N. A. Katulak, R. Wu, J. A. Dubin, A. Latimer, B. Meandzija, T. P. George, P. Jatlow, J. L. Cooney, et al. (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.

Trautmann, S. T. and G. van de Kuilen (2015). Belief Elicitation: A Horse Race among Truth Serums. *Economic Journal 125*(589), 2116–2135.

Tversky, A. and D. Kahneman (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.

Vogel, T. and M. Wanke (2016). *Attitudes and Attitude Change* (Second ed.). London and New York: Routledge.

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

WHO (2016). Pakistan health profile 2015. Technical report. Available at https://apps.who.int/iris/handle/10665/253769, last accessed January 26, 2019.

WHO (2018). Newborns: Reducing mortality. Available at http://www.who.int/news-room/fact-sheets/detail/newborns-reducing-mortality, last accessed September 10, 2018.

Wildschut, T., M. Bruder, S. Robertson, W. A. van Tilburg, and C. Sedikides (2014). Collective nostalgia: A group-level emotion that confers unique benefits on the group. *Journal of Personality and Social Psychology 107*(5), 844–863.

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

Winters, M. S., S. Dietrich, and M. Mahmud (2017). Perceptions of Foreign Aid Project Quality in Bangladesh. *Research and Politics 4*(4), 1–6.

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

World Bank (2015). *World Development Report 2015 – Mind, Society, and Behavior*. Washington, DC: World Bank.

World Health Organization (2018). Experiences with the WHO Safe Childbirth Checklist from around the world. Available at http://www.who.int/patientsafety/implementation/checklists/scc-photostory/en, last accessed October 3, 2018.

Young, A. (2017). Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results. Available at http://personal.lse.ac.uk/YoungA/ChannellingFisher.pdf, last accessed September 27, 2018.

Appendix.1 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 behavioural reactions (Tversky and Kahneman, 1981; Kahneman, 2003; Johnson and Goldstein, 2003; Hossain and List, 2012; Payne et al., 2013). The result is what is called the *framing effect.*[45] Stressing certain aspects invokes different associations and leads to different evaluations by the decision maker. Framing effects have been incorporated into theories on human behaviour 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."

[45]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: Most prominently and widely researched is the risky choice framing (risk of losing vs. risk of winning) as introduced by Kahneman and Tversky (1979). Attribute framing makes certain characteristics of a choice or good more salient (ground beef that is 75 percent lean vs. 25 percent fat). Lastly, goal framing where either punishment or reward is emphasized (Behavioral Science Solution, 2018). 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, e.g., with regard to smoking cessation, HIV screening as well as skin and breast cancer prevention (Kalichman and Coley, 1995; Detweiler et al., 1999; Schneider et al., 2001; Toll et al., 2007).

In order to abstract from the specific actors within our stetting, we named different actors (e.g., researchers, practitioners).[46] A qualitative investigation was conducted prior to the experiment to ensure that the correct terms were used to describe "local" versus "international" agents.[47] 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.[48] 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.

In addition to this traditional monetary outcome, we also collected measures suggested by other disciplines. Psychologists commonly assess the respondent's behaviour through time investments (Wildschut et al., 2014). Actual behaviour

[46]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.

[47]For this purpose, we talked to health-care 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 B.2 in the Appendix). There is a large consensus among respondents regarding the main countries associated with "international," namely Germany (24 percent), Malaysia (19 percent), USA (13 percent), Australia (8 percent). 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 2004's 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).

[48]After the debriefing, we offered participants to change their monetary contribution. 39 (16.5 percent) participants made use of this option. Generally, this led to an increase in contributions by on average one category (about 4,200 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.

measured by contributing money may be strongly influenced by general or situational economic living conditions of respondents. In case respondents face strong economic constraints, small or zero contributions might reflect a high neediness rather than lack of support for the intervention. Hence, we asked the participant's willingness to invest additional time to practice checklist use during regular working weeks. Further, in order to counter potential social desirability bias, we asked the participants to estimate the average monetary contribution of colleagues in other health facilities in the province. One's willingness to support an intervention might also be strongly determined by those beliefs about others' contribution. However, reporting one's perception about others might be subject to conformity bias, especially, in the Indonesian society, where a large focus is put on keeping one's face. Elicitation exercises based on introspection have been shown to reduce potential conformity bias in the experimental literature (Trautmann and van de Kuilen, 2015). Moreover, we use the outcome variable elicitation as a control variable in a further robustness test (see Appendix Table B.11). As expected, elicitation shows to be highly significant and positive, while the framing effect holds.

Appendix.2 Protocol

General Remarks[49]

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. 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.

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

Part A "Now, we would like to present you a new tool and would like to learn about your opinion towards 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 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*

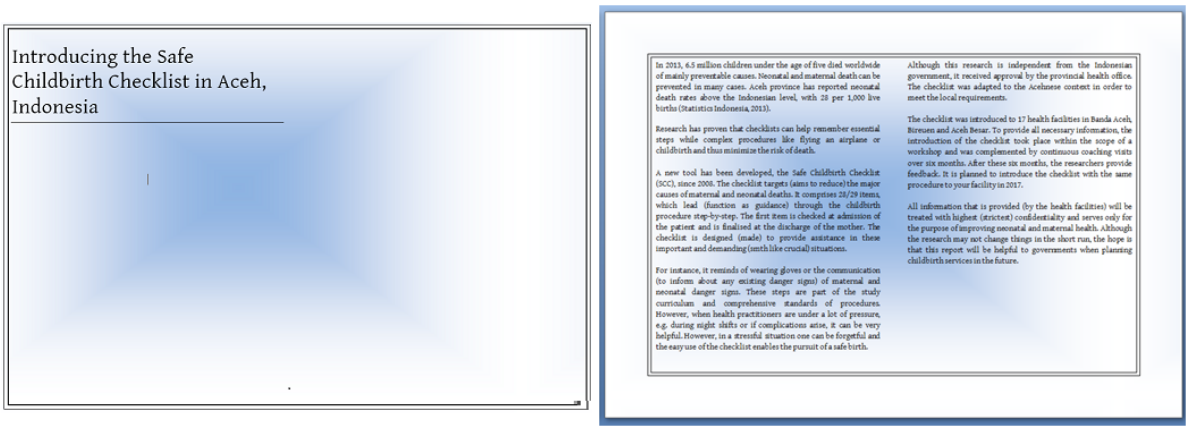
Signs" on checklist to the midwife or doctor] "All these steps are already part of the study curriculum. Hence, every checklist item is easy to understand. Generally, most of the health workers already practice these important steps in the delivery process. The checklist just has the purpose to remind you of all the important steps during the delivery process. Especially, when health practitioners are under a lot of pressure, e.g., 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 B.1 SCC Leaflet

Page 1 Page 2



Source: Authors' own depiction.

"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 5,000; 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; 5,000; 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 the Harvard School of Public Health and the World Health Organization. Local researchers from Syiah Kuala University 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 these information change your attitude towards 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.]*

Social desirability index We modify social desirability questions developed by Kemper et al. (2014) to reflect social desirability norms in the Acehnese context. The social desirability index was constructed by adding up the top categories (5 and 6) indicated in the subsequent questions.

Items

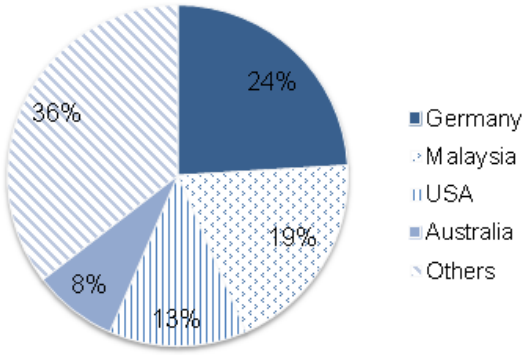
1. "In an argument, I always remain objective and not become emotional."
2. "Even if I am sad, I always smile when talking to others."
3. "When talking to someone older, I always listen carefully to what s/he says."
4. "When I had the chance to donate for religious purposes, I always contributed a lot."
5. "Sometimes I only help people if I hope to get something in return."

Answers

1. Disagree strongly
2. Disagree
3. Rather disagree
4. Rather agree
5. Agree
6. Agree strongly
7. Not applicable

Appendix.3 Figures

Figure B.2 Distribution of "International" Country Perceptions



Note: Based on "If you think of activities, programs or projects by internationals, which countries come first to your mind?"

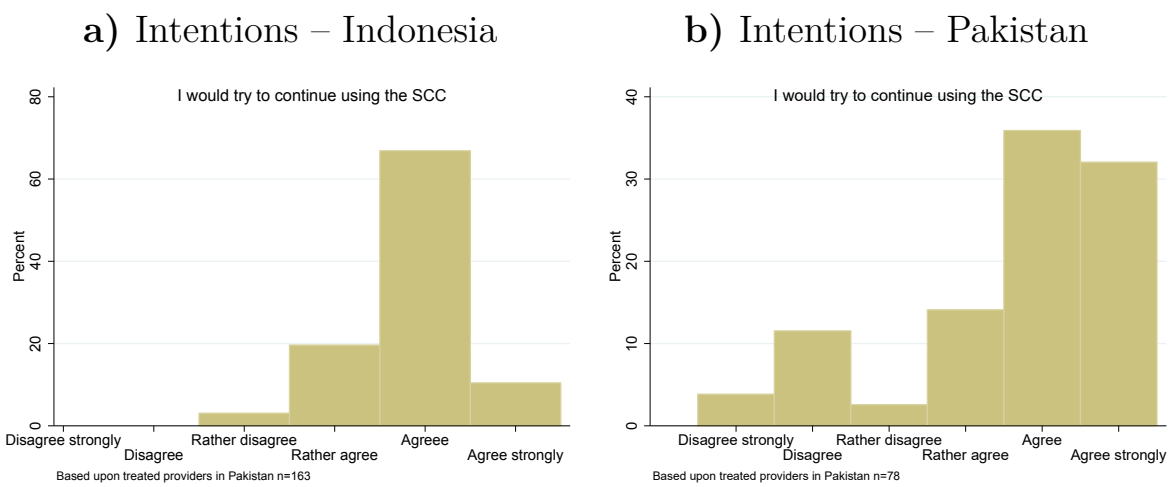
Source: Authors' depiction.

Descriptive Statistics

Corresponding to the high pre-intervention commitment, which we observed among midwives, there is a high level of reported intentions. Yet, Figures B.3a and B.3b indicate that there is some distinct variation within and across the settings.

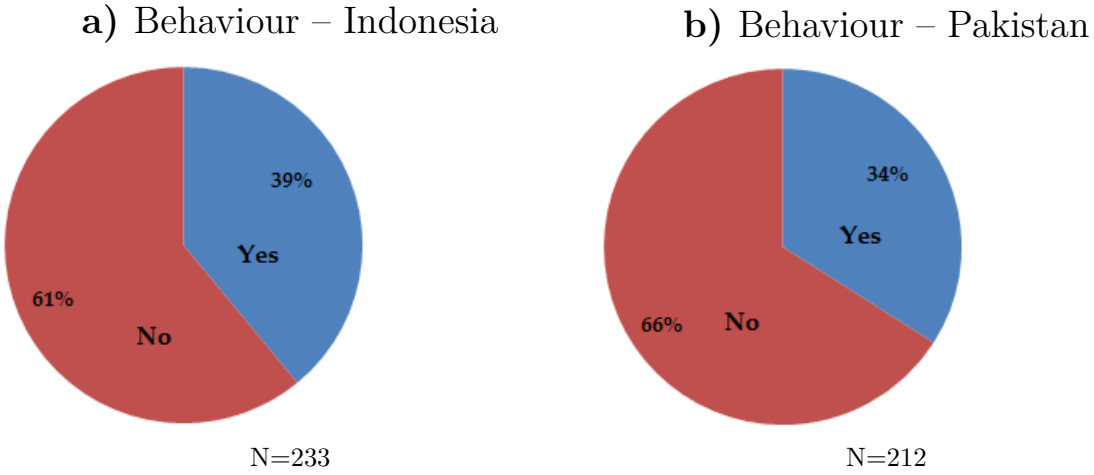
Yet, Figures B.4a and B.4b suggest a much lower level of actual uptake, which is examined in the regressions. Tables B.1 and B.2 provide a more comprehensive overview of descriptive statistics corresponding to Figures B.3a, B.3b, B.4a and B.4b.

Figure B.3 Intentions to use the Safe Childbirth Checklist



Source: Authors' calculation based on survey data.

Figure B.4 Actual use of the Safe Childbirth Checklist



N=233 N=212

Source: Authors' calculation based on clinical observations.

Table B.1 Summary Statistics for Indonesian data

	Full N	Full Max	Full Min	Full Mean	Full SD
Actual Behaviour:					
Active SCC Use	219	1	0	0.389	0.489
Intended Behaviour					
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 Centre	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

Table B.2 Summary Statistics for Pakistani data

	Full N	Full Max	Full Min	Full Mean	Full SD
Actual Behaviour:					
Active SCC Use	212	1	0	0.344	0.476
Intended Behaviour					
Would try to use SCC even if copies are not provided	78	6	1	4.628	1.452
Would recommend the SCC to fellow colleagues	78	6	1	5.141	1.090
Using the SCC in my professional role is	79	6	1	5.380	0.821
Ease to use SCC in work environment	79	6	1	4.962	1.305
SCC is supported by superiors	58	6	1	5.155	1.508
Urban (1) — Rural (2)	80	1	0	0.813	0.393
Open 24/7	80	1	0	0.150	0.359
Facility Type: Health Facility	80	1	0	0.2125	0.412
Facility Type: Community Midwife	80	1	0	0.5625	0.500
Facility Type: Lady Health Visitor	80	1	0	0.225	0.420
District: Haripur	80	1	0	0.450	0.501
District: Nowshera	80	1	0	0.550	0.501

Appendix.4 Analytical Appendix

Additional Results – Theory of Planned Behaviour

One concern might be that the variation in the outcome used for the analysis based on the TPB determinants (namely intentions) clusters strongly among high values. We, thus, apply in Table B.3 an alternative coding, where we recode the outcome as a binary variable, equalling one for the top categories "agree" (5) and "agree strongly" (6) and zero for the remaining categories. Results are qualitatively unchanged. Due to the limited number of clusters considered in the study, we also consider a standard error correction based on the method by Cameron et al. (2008). The results underline the pronounced role of subjective norms in Indonesia and the significant effect of perceived behavioural control in Pakistan. Results for actual SCC use become insignificant in Indonesia.

Finally, Table B.5 reruns the regressions from the main Tables 1 and 2 considering a principal component analysis based aggregate of all TPB determinants. While the TPB aggregate index in column 1b is not significantly correlated with intentions in Pakistan, when conditioning on control variables, remaining coefficients are positive and point to a positive direction. The positive correlations underline the explanatory power of the incentivizing factors considered.

Table B.3 TPB – Binary Outcome

	Intended SCC Use:			
	Pakistan		Indonesia	
	(1a)	(1b)	(2a)	(2b)
Attitudes:				
SCC in professional role: 1 "completely useless" – 6 "completely useful"				
	0.930***	0.704**	0.451***	0.317**
p-value	(0.007)	(0.025)	(0.006)	(0.013)
Subjective Norms:				
SCC is supported by superiors: 1 "not at all" – 6 "completely"				
	0.508	0.244	0.700***	0.444***
p-value	(0.118)	(0.475)	(0.009)	(0.003)
Perceived Behavioural Control:				
Ease of SCC in work environment: 1 "very difficult" – 6 "very easy"				
	0.763**	0.675**	0.303	-0.057
p-value	(0.011)	(0.041)	(0.166)	(0.746)
N	78	78	163	163
Control variables	No	Yes	No	Yes
Mean of dep. var.	4.628	4.628	4.847	4.847
Median of dep. var.	5	5	5	5
SD of dep. var.	1.452	1.452	0.634	0.634

Note: All regressions are based upon the treated providers. Adjusted regressions (b) additionally control for a variable indicating the facility type, a binary variable indicating rural/urban location, a variable indicating the district and for the Pakistani data a binary variable indicating whether the facility is open 24/7. Standard errors (SE) are clustered at the facility level. Asterisks indicate p-values according to: * p<0.1, **p<0.05, *** p<0.01.

Table B.4 TPB – Intentions and Behaviour: Wild Bootstrapped SE

	Intended SCC Use: Indonesia (1a)	Actual SCC Use: Pakistan (2a)	Actual SCC Use: Indonesia (2b)
Attitudes:			
SCC in professional role: 1 "completely useless" – 6 "completely useful"			
	0.454***	0.655***	-0.364
WB p-value	(0.004)	(0.000)	(0.505)
Subjective Norms:			
SCC is supported by superiors: 1 "not at all" – 6 "completely"			
	0.536*	0.207	0.642
WB p-value	(0.072)	(0.320)	(0.503)
Perceived Behavioural Control:			
Ease of SCC in work environment: 1 "very difficult" – 6 "very easy"			
	0.261	0.306***	0.038
WB p-value	(0.102)	(0.000)	(0.432)
N	163	212	218
Control variables	No	No	No
Mean of dep. var.	4.847	0.344	0.389
Median of dep. var.	5	–	–
SD of dep. var.	0.634	0.476	0.489

Note: Intended SCC Use was measured via the question "Would you try to use SCC even if copies are not provided anymore? (1 disagree strongly – 6 agree strongly)." Actual SCC Use was meassured via trained observers and is coded as a binary outcome variable. All regressions are based upon the treated providers. Standard errors (SE) are clustered at the facility level and wild cluster bootstrapped due to the small number of clusters (15 facilities), following Cameron et al. (2008). No bootstrapping is provided for intended SCC use in Pakistan as a sufficient number of clusters (70) was sampled. Asterisks indicate p-values according to:
* p<0.1, **p<0.05, *** p<0.01.

Table B.5 TPB – Intended SCC uptake: Principal Component Analysis

	Intended Behaviour			
	Would use SCC even if copies are not provided			
	1 "disagree strongly" – 6 "agree strongly"			
	Pakistan		Indonesia	
	(1a)	(1b)	(2a)	(2b)
PCA:				
Index based on the three TPB determinants				
	0.360*	0.174	0.232***	0.095**
p-value	(0.069)	(0.372)	(0.003)	(0.019)
N	57	57	219	219
Control variables	No	Yes	No	Yes
SD of dep. var.	1.452	1.452	0.634	0.634
	Actual Behaviour			
	Was SCC actively used or looked at during delivery?			
	0 "No" – 1 "Yes"			
PCA:				
Index based on the three TPB determinants				
	0.171***	0.088**	0.008	0.048***
p-value	(0.001)	(0.025)	(0.742)	(0.000)
N	212	212	219	219
Control variables	No	Yes	No	Yes

Note: All regressions are based upon the treated providers. The PCA treatment is constructed based on the first component of a principal component analysis on the three theory of planned behaviour determinants: attitudes, subjective norms and perceived behavioural control. Adjusted regressions (b) additionally control for a variable indicating the facility type, a binary variable indicating rural/urban location, a variable indicating the district and a binary variable indicating whether the facility is open 24/7. Standard errors (SE) are clustered at the facility level. Asterisks indicate p-values according to: * p<0.1, **p<0.05, *** p<0.01.

Additional Results – Framing Experiment

For the framing experiment, we find that the groups which were internationally or locally framed are generally balaned (both in the full and reduced sample as depicted in Tables B.6 and B.7). 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).

Table B.6 Experimental Balance – Full Sample

	Full N	Full Mean	Full SD	Control Mean	Control SD	Treat Mean	Treat SD	p-value difference
Facility Type	236	1.538	–	1.690	–	1.433	–	0.021**
Gender (1=m, 2=f)	236	2.000	–	2.000	–	2.000	–	–
Age (Years)	236	– 33.314	7.493	33.650	7.806	33.112	7.316	0.593
Education (Years)	236	15.051	0.527	15.020	0.603	15.067	0.462	0.619
Experience (Years)	236	9.576	7.271	9.690	7.736	9.537	6.979	0.886
Sufficient income	236	3.208	1.008	3.160	1.012	3.246	1.014	0.526
Financial problems	236	1.678	–	1.720	–	1.642	–	0.081*
Strategic donation	236	4.657	1.264	4.710	1.225	4.627	1.296	0.564
Social acc. Index	236	3.411	0.838	3.450	0.821	3.381	0.857	0.513
Social acc. # 1	236	4.966	0.690	5.000	0.778	4.940	0.622	0.480
Social acc. # 2	236	4.568	1.027	4.600	0.932	4.545	1.101	0.650
Social acc. # 3	236	5.343	0.558	5.310	0.506	5.366	0.595	0.172
Social acc. # 4	233	4.644	1.074	4.694	1.069	4.602	1.087	0.475
Social acc. # 5	236	2.229	1.254	2.250	1.298	2.216	1.235	0.784
Paperwork: too much	236	2.814	1.343	3.000	1.497	2.664	1.195	0.173
Routines ease work	236	5.153	0.734	5.150	0.626	5.179	0.764	0.660
Previous SCC experience	236	2.564	1.831	2.500	1.795	2.627	1.871	0.536
Previous SCC use	236	0.547	–	0.540	–	0.560	–	0.772
Access to resources	236	3.470	0.517	3.530	0.502	3.425	0.526	0.080*
Team effic. indicator	236	5.246	0.513	5.220	0.462	5.261	0.547	0.570
Part. in loc. projects	236	1.831	–	1.870	–	1.806	–	0.235
Part. in int. projects	236	1.898	–	1.880	–	1.910	–	0.511
Part. in donor projects	236	1.907	–	1.920	–	1.896	–	0.511

Note: Based upon the full sample with N denoting 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<0.1, **p<0.05, *** p<0.01.

Table B.7 Experimental Balance – Reduced Sample

	Full N	Full Mean	Full SD	Control Mean	Control SD	Treat Mean	Treat SD	p-value difference
Facility Type	170	1.500	–	1.618	–	1.409	–	0.050*
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	0.232
Education (Years)	170	14.994	0.516	14.974	0.565	15.011	0.478	0.742
Experience (Years)	170	8.888	7.094	8.974	7.494	8.849	6.824	0.908
Sufficient Income	170	3.200	1.069	3.118	1.083	3.269	1.065	0.348
Financial problems	170	1.741	–	1.763	–	1.720	–	0.396
Strategic donation	170	4.606	1.411	4.658	1.381	4.581	1.440	0.613
Social acc. Index	170	3.329	0.827	3.316	0.852	3.344	0.814	0.808
Social acc. # 1	170	5.000	0.738	4.987	0.887	5.011	0.599	0.834
Social acc. # 2	170	4.459	1.142	4.461	1.026	4.462	1.239	0.991
Social acc. # 3	170	5.429	0.584	5.408	0.521	5.452	0.634	0.436
Social acc. # 4	167	4.545	1.063	4.649	1.065	4.457	1.063	0.239
Social acc. # 5	170	2.118	1.286	2.184	1.334	2.065	1.258	0.375
Paperwork: too much	170	2.906	1.364	3.145	1.547	2.720	1.174	0.150
Routines ease work	170	5.100	0.727	5.079	0.648	5.151	0.722	0.471
Previous SCC experience	170	2.765	1.983	2.632	1.945	2.882	2.026	0.298
Previous SCC use	170	0.541		0.553	–	0.538	–	0.854
Access to resources	170	3.441	0.498	3.513	0.503	3.387	0.490	0.060*
Team effic. indicator	170	5.200	0.443	5.158	0.434	5.226	0.445	0.459
Part. in loc. projects	170	1.829	–	1.868	–	1.796	–	0.131
Part. in int. projects	170	1.918	–	1.895	–	1.935	–	0.272
Part. in donor projects	170	1.935	–	1.934	–	1.935	–	0.959

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<0.1, **p<0.05, *** p<0.01.

Table B.8 Framing Experiment – Wild Bootstrapped SE

Financial Contribution in support of SCC project (in IDR)		
	(a)	(b)
Framing: 1="internat."	557.624	1,283.772**
WB p-value	(0.404)	(0.032)
N	165	165
Control variables	no	Yes
Mean of dep. var.	4,757.576	4,757.576
SD of dep. var.	4,711.366	4,711.366

Note: See Table 3. Standard errors (SE) are clustered at the facility level and wild boostrapped 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<0.1, **p<0.05, *** p<0.01.

Table B.9 Framing Experiment – Covariates

	Recom-mendation	Time Investment	Own Contribution	Elicitation	PCA
Public Hospital	-0.063	-1.044	-3,444.525***	415.641	-0.710*
p-value	(0.595)	(0.073)	(0.0000)	(0.816)	(0.064)
WB p-value	(0.651)	(0.134)	(0.002)	(0.695)	(0.200)
Private Hospital	-0.217	0.826	-1,093.573	1,162.358	0.042
p-value	(0.296)	(0.265)	(0.667)	(0.337)	(0.923)
WB p-value	(0.302)	(0.344)	(0.541)	(0.454)	(0.873)
Social Acc. Index	0.132*	0.934***	825.220*	-81.462	0.446***
p-value	(0.071)	(0.000)	(0.091)	(0.704)	(0.002)
WB p-value	(0.082)	(0.000)	(0.114)	(0.637)	(0.000)
Paperwork: too much	-0.149***	-0.637***	-978.225***	-599.969**	-0.443***
p-value	(0.003)	(0.000)	(0.002)	(0.019)	(0.000)
WB p-value	(0.004)	(0.002)	(0.002)	(0.012)	(0.004)

Note: All specifications are based upon the sample limited to those respondents without prior SCC contact (refer to Table B.10). Community health clinics (puskesmas) constitute the comparison group regarding the facility type. 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<0.1, **p<0.05, *** p<0.01.

Our alternative outcome measures are first, whether respondents would recommend the SCC to fellow colleagues, second, whether they would be willing to invest additional

time for the SCC project, third, how high they estimate the average contribution by others (belief elicitation) and fourth an index of all four outcome measures, using principal component analysis (PCA). Estimates in Table B.10 show robustly positive coefficients, when controls are included and reach statistical significance for recommending the SCC to others and for the PCA-index. Here, however, the financial contribution is the variable that explains the major part of the variation in the index. Hence, our results suggest that the intervention is increasingly supported by the respondents, if it is perceived as an internationally-led endeavor.

When being financially incentivized to assess the potential answer of an anonymous third person (belief elicitation), opportunity costs of not revealing the own true assessment increase. We, thus, incentivized respondents with an additional pay-off of 10,000 IDR to estimate the average contribution category of respondents at other facilities. In a resource constrained setting the beliefs about the willingness of others to contribute could provide more accurate information about preferences as they are less subject to idiosyncratic financial situations of respondents. While those beliefs enter as hypothesized significantly in Table B.11, the framing remains independently significant.

Table B.10 Framing Experiment – Alternative Outcomes

	Recommendation 1–6		Time Investment 5 min. categories		Elicitation IDR		PCA All outcomes	
Framing: 1="internat."	0.049	0.126*	-0.151	0.095	605.929	769.956	0.108	0.317**
p-value	(0.535)	(0.058)	(0.404)	(0.624)	(0.447)	(0.304)	(0.525)	(0.012)
RI p-value	(0.600)	(0.122)	(0.668)	(0.746)	(0.342)	(0.239)	(0.584)	(0.0530)
WB p-value	(0.531)	(0.076)	(0.370)	(0.571)	(0.452)	(0.282)	(0.525)	(0.010)
N	167	167	167	167	167	167	167	167
Control variables	No	Yes	No	Yes	No	Yes	No	Yes
Mean of dep. var.	5.108	5.108	5.084	5.084	7,365.269	7,365.269	-0.117	-0.117
SD of dep. var.	0.581	0.581	2.237	2.237	3,950.536	3,950.536	1.289	1.289

Note: See Table 3. 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. (2016). Asterisks indicate p-values based on SE clustered at the facility level: *p<0.1, **p<0.05, *** p<0.01.

Table B.11 Framing Experiment – Elicitation as Control

Financial Contribution in support of SCC project (in IDR)	
Framing: 1="internat."	852.610*
p-value	(0.064)
Elicited Contribution of Others	0.5000***
p-value	(0.002)
N	165
Mean of dep. var.	4,757.576
SD of dep. var.	4,711.366

Note: See Table 3. Moreover, the elicited contribution of health practitioners from other facilities is added as a control variable. Standard errors (SE) are clustered at the facility level. Asterisks indicate p-values according to: *p<0.1, **p<0.05, *** p<0.01.

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 health care 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 to the SCC and induce, thus, a downward bias. Second, individuals with prior contact to the checklist might react heterogeneously due to more comprehensive information.

Full sample regression results controlling for prior contact, are shown in Table B.12 and are comparable to the findings presented in the main part. As a further robustness check we estimate a regression in Table B.13, 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 B.14. 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 heterogenous support for international and local actors, we also collected information on perceptions regarding local and international implementers. 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 B.15 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 B.15 supports this notion. Yet, in order to get some understanding of the channel variables, we consider some qualitative insights from open-ended questions in the main part.

Table B.12 Framing Experiment – Prior Contact as Control

	Recommendation		Time Investment		Own Contribution		Elicitation		PCA	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)	(5a)	(5b)
Framing: 1="internat."	0.058	0.128**	-0.048	0.177	537.557	1,206.299*	458.103	789.408	0.115	0.323***
p-value	(0.291)	(0.039)	(0.796)	(0.250)	(0.445)	(0.062)	(0.592)	(0.248)	(0.502)	(0.008)
WB p-value	(0.286)	(0.040)	(0.785)	(0.240)	(0.450)	(0.050)	(0.619)	(0.260)	(0.460)	(0.008)
N	230	230	230	230	226	226	230	230	226	226
Control variables	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: See Table 3. All specifications are based upon the full sample. 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<0.1, **p<0.05, *** p<0.01.

Table B.13 Framing Experiment – Interaction with Prior Contact

Financial Contribution in support of SCC project (in IDR)		
	(a)	(b)
Framing: 1="internat."	557.624	1,164.830**
p-value	(0.395)	(0.033)
Prior Contact × Local Framing	225.973	627.961
p-value	(0.835)	(0.547)
Prior Contact × International Framing	706.522	1,955.229
p-value	(0.547)	(0.105)
N	226	226
Control variables	No	Yes
Mean of dep. var.	4,757.576	4,757.576
SD of dep. var.	4,711.366	4,711.366

Note: See Table 3. The base category is No Prior Contact and Local Framing. Asterisks indicate p-values based on standard errors clustered at the facility level: * p<0.1, **p<0.05, *** p<0.01.

Table B.14 Framing Experiment – Ordered Probit Results

	Recommendation		Time Investment		Own Contribution		Elicitation	
Framing: 1="internat."	0.191	0.522***	-0.087	0.060	0.081	0.306**	0.129	0.188
p-value	(0.316)	(0.010)	(0.239)	(0.647)	(0.600)	(0.017)	(0.535)	(0.341)
N	167	167	167	167	165	165	167	167
Control variables	No	Yes	No	Yes	No	Yes	No	Yes

Note: See Table 3. 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<0.1, **p<0.05, *** p<0.01.

Table B.15 Framing Experiment – Association with Potential Channel Variables

	Control Capabilities	Implementation Skills	Funding Capabilities	Account-ability	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	(0.002)	(0.003)	(0.007)	(0.090)	(0.393)	(0.638)	(0.257)
WB p-value	(0.004)	(0.008)	(0.008)	(0.118)	(0.374)	(0.719)	(0.224)
N	230	230	230	230	230	230	230

Note: All specifications are based upon 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<0.1, **p<0.05, *** p<0.01.

Point Estimates – Previours Experience Table B.16 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. While the randomization ensured that the framing could be considered as exogenous, project participation is potentially endogenous regarding other traits of the surveyed respondent. However, as recent research by Nizalova and Murtazashvili (2016) and Bun and Harrison (2018) indicates, the interaction of an exogenous and an endogenous variable can be considered as exogenous, when controlling for the endogenous variable.[50] Moreover, balancing tests provided in Table 3 and B.16 underscore that previous participation is balanced across both framing treatments. The results in columns (1a-b) 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. 6,500-8,500 IDR (e.g., 0.45-0.65 US$) higher if framed international. Thus, the effect of the *attitude* towards the intervention in the unadjusted and adjusted specification is significantly higher if respondents knowing both implementers are framed internationally (p-value: 0.025 and 0.000, respectively). Respondents who previously participated in local projects do not contribute different amounts of money when faced with an international framing. However, 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, both significant with and without adjusting for controls (p-value: 0.012 and 0.052, respectively). Finally, row VII does not depict any

[50]Nonetheless, 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.

significant framing effects, if respondents did not have any prior experience. 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 B.16 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	2,708.333	4,202.892**
p-value	(0.237)	(0.019)
(II.) International Framing (0) × Int. participation (1) × Loc. Participation (1)		
β	-3,791.667***	-4,313.226***
p-value	(0.007)	(0.000)
Coefficient Equality Row (I) & (II)	0.025	0.001
(III.) International Framing (1) × Int. participation (0) × Loc. Participation (1)		
β	-2,291.667*	-1,196.631
p-value	(0.068)	(0.287)
(IV.) International Framing (0) × Int. participation (0) × Loc. Participation (1)		
β	-148.810	-537.176
p-value	(0.918)	(0.762)
Coefficient Equality Row (III) & (IV)	0.186	0.660
(V.) International Framing (1) × Int. participation (1) × Loc. Participation (0)		
β	-625.000	1,433.060
p-value	(0.710)	(0.507)
(IV.) International Framing (0) × Int. participation (1) × Loc. Participation (0)		
β	-4,791.667***	-4,184.609
p-value	(0.000)	(0.130)
Coefficient Equality Row (V) & (VI)	0.012	0.052
(VII.) International Framing (1) × Int. participation (0) × Loc. Participation (0)		
β	646.930	1,009.864
p-value	(0.463)	(0.200)
N	165	165
Control variables	No	Yes

Note: See Table 3. Standard errors (SE) are clustered at the facility level. Asterisks indicate p-values according to: *p<0.1, **p<0.05, *** p<0.01.