Benefit Salience and Cognitive Barriers to Programme Take-up: A Field Experiment

Submitted to *Behavioural Public Policy*

April 28, 2026

Abstract

Low take-up of welfare-improving programmes is often attributed to liquidity constraints, yet participation may lag even when monetary costs are negligible. We isolate the role of cognitive barriers and trust frictions in the context of China's Rural Toilet Revolution, a unique setting where full government funding effectively rules out the inability to pay channel. To disentangle the content of the signal from its source, we implemented a 3×2 factorial RCT that cross-randomises three information frames (benefits, harms, or technical details) against the presence of a peer messenger. We document a sharp divergence between stated intentions and revealed preferences. While all interventions uniformly increased short-run self-reported willingness to participate, actual administrative enrolment responded significantly only to benefit-salience messages. We detect no effect of peer messengers. Heterogeneity analyses show that enrolment effects are concentrated among villagers with higher baseline programme literacy, suggesting a complementarity between information and prior knowledge. These results imply that even when financial constraints are removed, the inability to quantify invisible returns remains a binding constraint, and effective policy must resolve this valuation gap to trigger adoption.

Keywords: Public Programme Take-up; Cognitive Barriers; Information Nudges; Sanitation; Field Experiment

JEL Codes: C93; D83; D91; I15; O12

1 Introduction

Governments globally invest heavily in welfare-improving programmes, yet take-up sometimes remains puzzlingly low. While traditional economic explanations focus on liquidity constraints or transaction costs, participation can lag even when monetary costs are negligible (Currie, 2004). This persistence points to a fundamental cognitive barrier: beneficiaries often struggle to perceive or quantify the invisible returns of adoption, such as long-term health improvements or environmental externalities, leading to a valuation gap that impedes action (Bai, Chi, Liu, Tang, & Xu, 2021; Ko & Moffitt, 2022). Behavioral science offers a potential solution: using information nudges to make these invisible benefits salient (Kremer, Rao, & Schilbach, 2019). Within this framework, Prospect Theory provides a dominant blueprint for design, positing that individuals are more motivated by avoiding losses than by acquiring equivalent gains (Kahneman, Tversky, et al., 1979). Consequently, a preferred policy prescription is to frame non-participation as a loss (e.g., highlighting health risks) to leverage loss aversion. However, does this principle hold universally in developing contexts where the status quo itself is dire?

In this paper, we isolate the role of cognitive barriers by exploiting the unique zero-cost setting of China's Rural Toilet Revolution, where government subsidies fully cover financial costs, ruling out liquidity constraints. We conducted a randomised controlled trial (RCT) in southwest China within the context of Rural Toilet Revolution (RTR), a massive state-led campaign to modernise sanitation infrastructure. Crucially, while the government fully funds hardware and installation, rendering the financial cost negligible, the returns to adoption, such as long-term health improvements, remain largely invisible and difficult for households to quantify. This distinct combination of negligible financial costs and non-salient benefits offers a natural laboratory to test whether the provision of information targeting cognitive barriers can drive participation in the absence of liquidity constraints. Additionally, we investigate the role of social proximity through the presence of a peer villager, because extant literature shows that even when accurate information is available, households may discount it if the source lacks social proximity or perceived credibility (Alsan & Wanamaker, 2018; BenYishay & Mobarak, 2014).

We implemented a 3×2 factorial design, with two design dimensions (information salience and messenger effect), which unbundles the content of the signal from its source. Along the

information salience dimension, we address the cognitive barriers by randomising the specific attribute highlighted in advocacy videos: (1) gain-framed benefits (health and environmental returns), (2) loss-framed benefits (harms of poor sanitation),[1] or (3) technical guidance (construction details). Along the messenger dimension, we address potential trust friction through the presence of a peer advocate, a local villager, against no such presence. This design allows us to distinguish the two channels: if the barrier is a failure to value invisible returns, benefit salience should dominate; if the barrier is trust friction, peer endorsement should drive adoption. We recruited 985 households in 93 villages with historically low coverage of sanitary toilets and measured two outcomes: immediate self-reported willingness to participate (WTP) and verified administrative enrolment ten months later.

Our results highlight a sharp divergence between stated intentions and revealed preferences. In the short run, we observe a substantial increase in self-reported WTP across all treatment arms; however, this increase is statistically indistinguishable from an active placebo group that viewed generic rural revitalisation content. This suggests that stated WTP in this context largely reflects survey priming or experimenter demand effects rather than genuine participation intention. In contrast, our long-run administrative data reveal economically meaningful treatment effects on actual enrolment. The gain-framed benefit-salience video significantly increases verified enrolment by 8.8 percentage points. When pooling gain- and loss-framed interventions, the 'benefit salience' treatment yields a combined 6.8 percentage point increase in take-up. Heterogeneity analyses indicate that these effects are concentrated among individuals with higher baseline programme literacy, suggesting that informational nudges act as a complement to foundational knowledge. Conversely, we detect no significant messenger effect, which suggests that when the primary barrier is a valuation problem, low-intensity social signals are insufficient to alter behaviour without substantive information on returns.

This study makes four primary contributions to the literature. First, we provide direct empirical evidence that cognitive barriers (inability to process benefits) act as an independent and

[1] This design arm is motivated by Prospect Theory, which posits that individuals exhibit loss aversion, weighing losses more heavily than equivalent gains (Kahneman et al., 1979). In the context of sanitation, this framing mirrors the Community-Led Total Sanitation (CLTS) approach, which successfully drives behavioural change by making the negative externalities of open defecation (e.g., disgust, disease transmission) salient, rather than solely emphasising health improvements (Kar & Pasteur, 2005).

binding constraint. While the development literature has extensively examined liquidity constraints (A. V. Banerjee, Banerji, Duflo, Glennerster, & Khemani, 2010; Ko & Moffitt, 2022), identifying the specific role of cognitive frictions has proven difficult where monetary costs confound decision-making (Handel & Schwartzstein, 2018). By exploiting the RTR's zero-cost setting, we show that even when the price is zero, adoption is not automatic; resolving the valuation gap regarding invisible returns is a prerequisite for translating eligibility into enrolment.

Second, we contribute to the Water, Sanitation, and Hygiene (WASH) literature by refining the debate between pricing and education (Cohen & Dupas, 2010; Guiteras, Levinsohn, & Mobarak, 2015; Stockman et al., 2007). We demonstrate that the last mile of adoption for one-off durable assets differs from recurrent behaviours like handwashing. In this context, generic information is insufficient; interventions must specifically simplify the cost-benefit calculus of non-salient future returns to be effective (Loyalka et al., 2013; Manoli & Turner, 2014).

Third, we document the methodological importance of distinguishing between stated and revealed preferences in informational experiments. The contrast between our null results relative to placebo in survey data and the robust effects in administrative data underscores the limitations of WTP as a proxy for behaviour in policy settings. Our findings suggest that genuine belief updating capable of driving costly behavioural change occurred only when the specific cognitive barrier was addressed, a pattern that self-reported survey responses might fail to capture.

Furthermore, our results offers new insights to behavioral public policy and on the application of Prospect Theory in developing contexts. While loss aversion typically suggests that framing outcomes as 'losses' should be more motivating than equivalent 'gains', we find the opposite: highlighting the health harms of the status quo failed to drive enrolment. We attribute this to hedonic adaptation: rural households may have normalised the negative externalities of unimproved sanitation (e.g., odour, flies) as their reference point. Consequently, messages emphasising these conditions are not processed as a 'loss' from a neutral baseline, but rather as a description of the status quo. This finding provides a crucial nuance to the design of informational nudges: effective framing requires calibrating the message to the target population's

subjective reference point.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background. Section 3 presents the conceptual framework. Section 4 details the experimental design and data. Section 5 outlines the empirical strategy. Section 6 presents the main results. Section 7 concludes.

2 Institutional background

Poverty eradication has long been China's top development priority. By 2021, the country had successfully lifted over 800 million citizens out of poverty, defining a new baseline for global development (World Bank, 2022). Following this achievement, the national strategy has shifted toward 'Rural Revitalisation' and 'Common Prosperity', aiming to reduce the persistent urban-rural divide in living standards and public services (Brookings Institution, 2022). The Rural Toilet Revolution (RTR) is a cornerstone of this agenda. Unlike early-stage infrastructure projects that focused solely on access, the current phase of the RTR targets the quality of sanitation, especially the adoption of harmless waste treatment technologies, to improve public health and ecological outcomes (Cheng et al., 2018; Li et al., 2021; Liu, Zang, & Yang, 2020). Consequently, the policy focus has moved from supply-side construction to demand-side adoption, making household take-up the primary metric of success.

We conducted our RCT in six townships within a southwestern municipality of China. This region was selected for three strategic reasons. First, it presents a substantial urban-rural quality gap in sanitation. This discrepancy implies significant scope for improvement and suggests that while households are accustomed to basic latrines, they may lack understanding regarding the benefits of advanced waste treatment. Second, the region represents a typical top-down implementation environment where the local government aims to reach a full coverage of sanitary toilet facilities and septic systems, relying heavily on frontline cadres to persuade villagers. Third, the demographic profile minimises potential cultural confounds. The study area has a population of over 1.1 million (37.5% rural) and is predominantly Han Chinese ($> 95\%$). Unlike regions where specific religious or ethnic customs significantly dictate hygiene practices

(Geruso & Spears, 2018), this cultural homogeneity helps ensure that resistance to adoption is not driven by unobserved cultural taboos regarding sanitation.

Most crucially, a defining feature of the RTR in our study context is its financing structure, which creates a sharp dichotomy between the visible costs of adoption and its invisible returns. To rapidly achieve coverage targets, the local government adopted a turnkey approach for two primary interventions: (1) sanitary toilet retrofitting (e.g., installing flush systems) and (2) septic system construction (specifically, three-chamber tanks for harmless waste treatment). Under this policy, the government covers the full cost of hardware, materials, and installation labour. Whether for a simple retrofit or a complex septic installation, the direct pecuniary cost to participating households is effectively zero.[2]

However, while the financial costs are fully removed, the benefits of adoption remain largely invisible. Unlike the tangible hardware or cash transfers provided by the government, the returns to sanitation, such as the interruption of faecal-oral disease transmission and long-term groundwater protection, are abstract, delayed, and difficult for households to observe directly. Despite the full subsidy, participation remains strictly voluntary: households must explicitly agree to the renovation. This requirement makes the adoption decision hinge on two non-monetary factors: perceiving the utility of participation (valuation) and trusting the credibility of the mandate (trust). This unique institutional setup provides an ideal setting to isolate these specific frictions since the 'inability to pay' is institutionally ruled out.

3 Conceptual Framework

To guide our experimental design, we integrate the classic economic taxonomy of take-up barriers with recent insights from behavioural science. Specifically, we draw on three distinct strands of literature, namely salience (addressing cognitive barriers), reference-dependent preferences (framing), and messenger effects (addressing relatability issues, to propose our core hypotheses.

[2]In the current phase, households can choose between two modes: (1) a government-organised 'turnkey' service; or (2) a self-build model, where households are reimbursed a lump sum up to 2,000 RMB (approx. US$ 280) upon passing inspection. Crucially, this subsidy is calibrated to fully cover local market rates for materials and labour. Thus, regardless of the mode, the net monetary cost is non-positive, ensuring that liquidity constraints do not bind.

Cognitive Barriers and Salience Standard economic models typically attribute incomplete take-up to low net monetary utility, while a large body of literature identifies non-monetary frictions as binding constraints. Currie (2004) categorises these barriers into stigma and transaction costs, explicitly highlighting the lack of information as a critical cost of claiming and showing that the burden of learning about a programme acts as a de facto entry barrier. Recent research refines this notion into a broader framework of “cognitive barriers" (Ko & Moffitt, 2022), arguing that the complexity of application procedures imposes a “learning cost" as a cognitive tax, that discourages participation even when eligibility is clear.

Crucially, these cognitive barriers are amplified by liquidity constraints and economic precarity. As Mullainathan and Shafir (2013) demonstrate, poverty imposes a pervasive “bandwidth tax” on the mind. The constant preoccupation with managing scarce resources creates “tunnelling”, which refers to a focus on immediate necessities that depletes the executive control required for complex, long-term planning. Empirical evidence from Lichand and Mani (2020) confirms this mechanism by showing that income shocks directly divert attention and impede cognitive function among farmers in drought-prone Brazil, thereby reducing decision-making quality.

How can policy interventions address this well-documented cognitive barrier? A rich literature suggests that making benefits salient is a potent mechanism. Chetty, Looney, and Kroft (2009) demonstrate that agents often optimise based on salient attributes rather than full information. Consequently, when benefits are opaque or delayed, common features of preventive health investments, individuals systematically undervalue them (Dupas, 2011). This valuation gap can be bridged by informational interventions that heighten visibility. For instance, Jensen (2010) find that informing students of the true returns to schooling boosts enrolment by correcting subjective beliefs, while Bhargava and Manoli (2015) show that simplifying benefit notices to make potential gains salient effectively reduces psychological frictions for low-income households.

Therefore, in our zero-cost setting, we posit that the binding constraint is not financial, but cognitive. Villagers operating under liquidity constraints face cognitive barriers that prevent them from processing the invisible returns of sanitation. An informational nudge that makes

these attributes salient should function as a critical cognitive prosthetic.

- **Hypothesis 1 (Salience Effect):** Providing salient information about the programme's value (regardless of framing) will increase participation relative to a generic placebo.

Information Framing If salient information is required to overcome cognitive barriers, how should this information be structured? The foundational work of **?** challenges expected utility theory by positing that individuals evaluate outcomes as deviations from a reference point. Crucially, the value function is asymmetrical: generally steeper for losses than for gains. This implies that the disutility of a loss looms larger than the utility of an equivalent gain.

This asymmetry informs the goal framing effect in policy communication. As systematised by Levin, Schneider, and Gaeth (1998), when the goal is to persuade individuals to adopt a specific behaviour, messages emphasising the negative consequences of inaction (loss frame) are typically more effective than those highlighting positive outcomes (gain frame). Empirical evidence validates the robustness of these principles across diverse contexts (Homar & Cvelbar, 2021; Ruggeri et al., 2020). For instance, Ganzach and Karsahi (1995) find that credit card usage messages framed as losses were more than twice as effective as gain-framed alternatives, demonstrating that leveraging loss aversion is a powerful tool for overcoming inertia.

A straightforward application of these behavioural principles yields our standard framing hypothesis:

- **Hypothesis 2a (Standard Loss Aversion):** Information nudges framed as *losses* (emphasising the harms of the status quo) will be more effective in driving enrolment than those framed as *gains*.

However, this prediction may not hold in the context of chronic deprivation. A critical boundary condition for loss aversion is the location of the reference point. Research on hedonic adaptation suggests that individuals habituate to stable circumstances, causing their internal reference points to align with their current reality (Brickman, Coates, & Janoff-Bulman, 1978; Frederick & Loewenstein, 1999). In the context of rural sanitation, households may have lived with traditional pit latrines for decades. Through the lens of hedonic adaptation, the negative

externalities of these latrines (e.g., odour, flies) are no longer processed as deviations from a healthy norm, but rather as the neutral status quo. In this scenario, a loss-framed message fails to trigger aversion because the audience does not perceive these conditions as a "loss" relative to their adapted baseline. Conversely, a gain-framed message (visualising a modern, tiled toilet) projects a tangible deviation above the reference point, creating an aspirational gap that motivates action.

- **Hypothesis 2b (Hedonic Adaptation):** Contrary to standard loss aversion, gain-framed messages will be more effective than loss-framed messages, as they project a positive deviation from a normalised, low-quality status quo.

Social Proximity Even if information is salient and framed effectively, a final barrier remains: social proximity. In developing contexts, households may disregard objective information if the source is perceived as lacking relatability. Empirical evidence in development economics highlights that social proximity is a key determinant of trust. For instance, BenYishay and Mobarak (2014) demonstrate that peer farmers are significantly more effective at inducing technology adoption than government extension workers, as they are perceived as sharing similar incentives and constraints. Similarly, in the health domain, while Alsan and Wanamaker (2018) show that the historical erosion of trust (e.g., Tuskegee) leads to sharp declines in healthcare utilisation, Alsan, Garrick, and Graziani (2019) show the converse: reducing the social distance between provider and patient (e.g., via race-concordant messengers) can resolve these frictions and significantly boost adoption. Together, these studies suggest that the messenger's identity dictates cognitive and psychological receptivity. In rural sanitation, villagers may harbour scepticism toward top-down mandates. A local peer messenger serves as a mechanism of social proximity, signalling that the programme is safe, feasible, and beneficial for people like us.

- **Hypothesis 3 (Messenger Effect):** Information delivered by a peer villager (high social proximity) will be more effective in driving enrolment than information delivered by a generic narrator, by bridging the resonance gap and mitigating trust frictions.

4 Experimental design

This section describes the experimental framework, including the treatment design, recruitment strategy, randomisation, implementation timeline, and outcome measurement. We implement a village-level cluster-randomised 3×2 factorial design plus an active-placebo control.

4.1 Design of treatments

The main intervention invited villagers to watch a short video introducing China's RTR and the broader rural revitalisation strategy. In the videos, we focus on three core attributes: (1) the benefits (loss- or gain-framed) of sanitary versus unhygienic living conditions; (2) technical details on the construction, use, and maintenance of relevant sanitary facilities; and (3) the purpose and broader policy rationale behind the RTR. The benefits and harms components convey closely related information but vary the framing (gains versus losses). Regarding technical requirements, we first illustrate the construction process of a standard sanitary toilet and three-chamber septic systems, and then provide guidance on daily use, management, and routine maintenance of the upgraded sanitation facilities.

Building on this content, we implement a two-dimensional treatment design: an *information salience* dimension and a *messenger presence* dimension. Along the *information salience* dimension, we make one of three attributes—benefits (BE), harms (HA), or techniques (TE)—more salient by allocating more time and more explicit explanations to that focal attribute, while holding total video duration constant. Along the *messenger presence* dimension, we vary whether a peer villager who previously participated in the RTR appears in (part of) the video to deliver the call-for-participation message.[3]

Figure 1 summarises the basic structure shared by all experimental videos. Each video contains three parts: (1) an introduction to the RTR and the rural revitalisation strategy, including the call-for-participation message; (2) an attribute module covering technical attributes

[3]The peer villager selected as the messenger was intended to maximise audience resonance through representativeness and perceived credibility. First, he is male, which aligns with the fact that most household decision-makers in this region are men. Second, his age (approximately 55 years old) is near the average age in the covered towns, and he speaks only the local dialect, a common characteristic in this region. Finally, we avoided choosing an individual with highly specific or atypical characteristics (e.g., exceptional appearance or uncommon affluence) to enhance relatability.

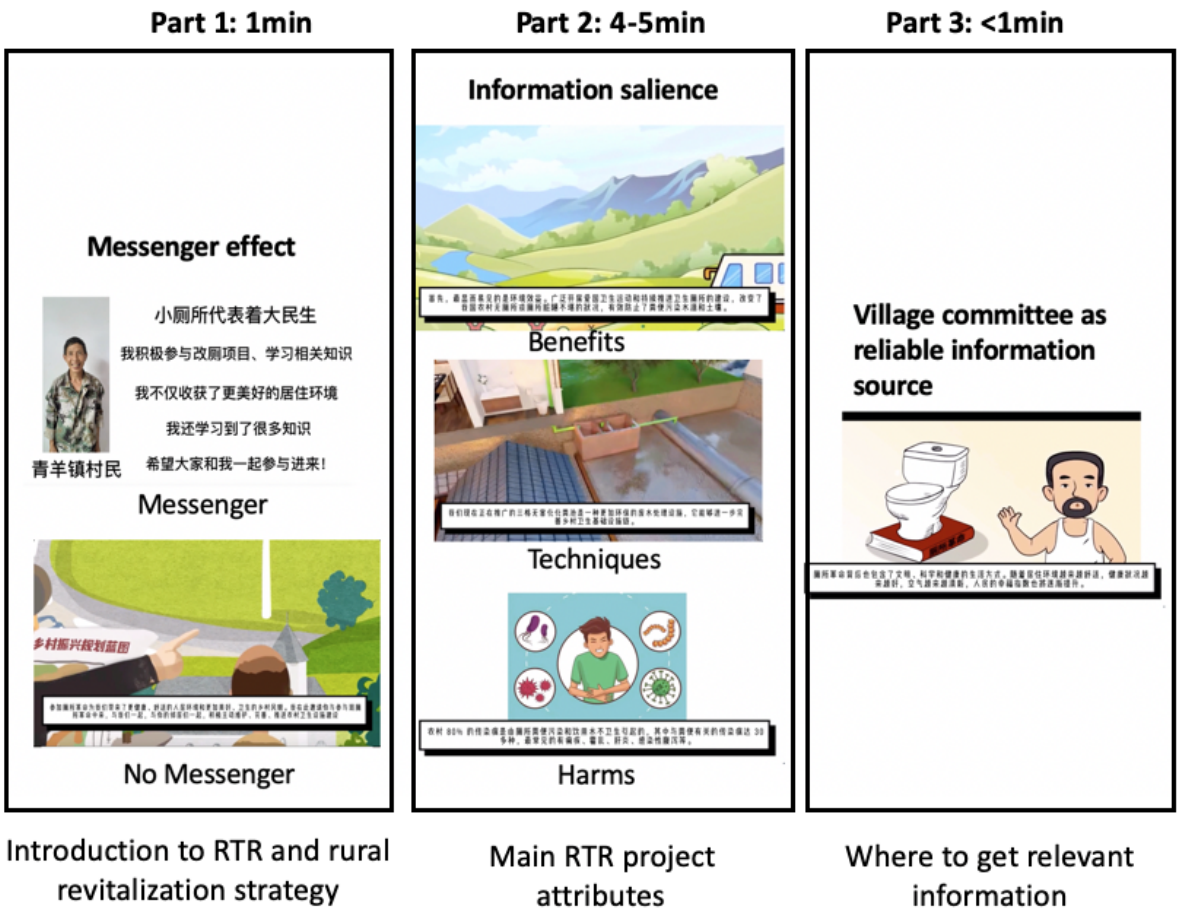

Figure 1: Video structure

of participation, benefits of RTR, and harms of poor sanitation, with one attribute emphasised depending on the treatment arm; and (3) closing remarks that direct viewers to the village committee as a reliable source of programme information.

Table 1 presents the 3×2 factorial design (six treatment arms). In addition, the control group was assigned an active placebo: a video introducing broad aspects of China's rural revitalisation strategy, with a specific focus on agricultural modernisation and economic development. To minimise implicit sanitation priming, we selected a placebo video that excludes imagery related to sanitation infrastructure, latrines, or domestic hygiene.

4.2 Randomisation

To study the impact of the intervention in settings with meaningful scope for adoption, we selected six townships with the lowest coverage rate of sanitary toilets among the 27 sub-district offices/townships in the region. To reduce risks of spillovers (e.g., information sharing across

Table 1: Three-by-two factorial design

Information Salience	**Messenger Presence**	
	Villager presence (ME)	No villager presence (NO-ME)
Benefits (BE)	• 2 min introduction to pecuniary and non-pecuniary benefits of participation in detail • 30s–50s for HA and TE • With the presence of a peer villager	• 2 min introduction to pecuniary and non-pecuniary benefits of participation in detail • 30s–50s for HA and TE • Without the presence of a peer villager
Harms (HA)	• 2 min introduction to harms caused by poor sanitation in detail • 30s–50s for BE and TE • With the presence of a peer villager	• 2 min introduction to harms caused by poor sanitation in detail • 30s–50s for BE and TE • Without the presence of a peer villager
Techniques (TE)	• 2 min introduction to how relevant facilities are built and maintained • 30s–50s for BE and HA • With the presence of a peer villager	• 2 min introduction to how relevant facilities are built and maintained • 30s–50s for BE and HA • Without the presence of a peer villager

experimental arms) and participant crossover, we implemented cluster-level randomisation at the village level. Specifically, we randomly assigned 93 villages in the six townships to one of the six factorial treatment arms or to the active-placebo control group. Within each village, we recruited 10–15 subjects and obtained 985 valid responses. Participation was entirely voluntary. To ensure that respondents did not come from the same household, we verified household identifiers using administrative records after registration. Figure 2 summarises the experimental workflow, and village-level summary statistics are reported in Table A2 (Appendix A). We set up the experiment site in the office of each village committee and recruited subjects with the support of village committees and township-level governments.

4.3 Intervention timing and data measurement

We implemented the video interventions between July and September 2022 across the study townships. Data collection proceeded in two phases, combining survey responses with administrative records.

Because information interventions can fail if participants do not attend to or retain the content, we first examine impacts on RTR-related knowledge. We focus on two knowledge outcomes: (1) the level of RTR understanding; and (2) whether the intervention leads to a

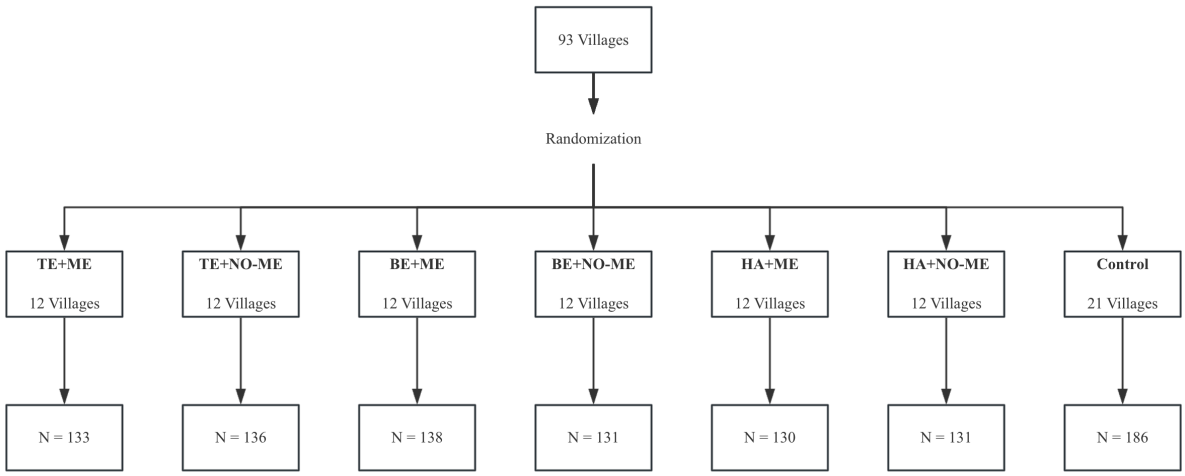

Figure 2: Randomisation

more complete understanding of RTR. To measure the level of RTR understanding, we use the respondent's selected definition of RTR, elicited by a survey question asking what RTR means. Responses are coded from the shallowest understanding (1) to the most complete (3).

For outcomes associated with RTR engagement, we consider (1) willingness to participate and (2) realised enrolment several months later. 'Willingness to participate' is the self-reported likelihood of signing up for the upgrade programme, measured on a 0–10 Likert-type scale from 0 (no participation for sure) to 10 (will participate for sure). This outcome is measured immediately after the intervention using an endline survey administered at the experiment site. 'Enrolment decision' is measured ten months post-intervention via telephone interviews and verified using administrative records. It is an indicator variable equal to 1 if the subject registered for the 2023 RTR project and 0 otherwise.

5 Empirical strategy

The primary objective of our empirical analysis is to test whether reducing cognitive barrier and trust friction through informational nudges increases participation in RTR. We proceed in three steps. First, to validate the 'first stage' of our intervention, we examine whether the treatments effectively improved participants' knowledge and understanding of the programme. Second, we analyse stated preferences by examining short-run changes in self-reported willingness to

participate immediately following the intervention. Finally, and most importantly, we estimate the treatment effects on revealed preferences using verified long-run enrolment data.

To estimate the Intention-to-Treat (ITT) effects of the informational nudges, we employ the following benchmark specification:

$$y_{ivp1} = \alpha_0 + T_{vp}{}'\beta + X_{ivp0}{}'\gamma + V_{vp0}{}'\lambda + \delta_p + \varepsilon_{ivp1} \tag{1}$$

where y_{ivp1} denotes the outcome of interest for individual i in village v and township p at endline. T'_{vp} is a vector of indicators representing the village-level assignment to the treatment arms, with the active-placebo group serving as the omitted reference category. The coefficient vector β captures the causal effect of the specific informational content (e.g., benefit salience, technical guidance) relative to generic policy attention. X'_{ivp0} is a vector of baseline individual controls, including demographics, prior RTR experience, baseline knowledge, and crucially, the baseline value of the outcome variable (y_{ivp0}). V'_{vp0} represents a set of baseline village-level covariates, including village area, average disposable income, gender ratio, poverty rate, and average household size, included to account for any chance imbalances in socio-economic conditions. δ_p denotes township fixed effects, absorbing time-invariant characteristics common to all villages within the same administrative unit. Standard errors are clustered at the village level, the unit of randomisation.

To improve precision and to capture changes net of baseline differences, we include the baseline level of willingness to participate, previous RTR experiences, and other baseline individual characteristics as controls. We also account for potential confounders from individual, village, and town features by incorporating baseline individual controls and baseline village controls.[4] We add town fixed effects δ_p to control for town characteristics that are common to individuals and villages within the same town.

To increase statistical power and map the analysis to the two-dimensional intervention design, we also estimate specifications that pool the six treatment arms along (1) the information-salience dimension (Technical vs Benefit vs Harm attributes) and (2) the messenger dimension

[4]To still account for potential differences among villages, we select a set of village level variables as proxies for villages' socio-economic conditions. The village-level covariates are areas, average disposable income, gender ratio, per cent of poor households, and average household size.

(messenger presence vs no messenger presence), effectively averaging over the other dimension. These pooled ITT estimates correspond to the results reported by design dimension in Section 6.

Knowledge outcomes (short-run, self-reported)

For knowledge outcomes, we estimate equation (1) for the indicator variable that equals 1 if the subject's reported RTR understanding at endline is more complete than at baseline, and 0 otherwise. For the level of understanding (an ordinal outcome), we estimate an ordered logit model of the following form:

$$Knowledge^*_{ivp1} = \alpha + T_{vp}{}'\beta + X_{ivp0}{}'\gamma + V_{vp0}{}'\lambda + \delta_p + u_{ivp1}, \tag{2}$$

$$Knowledge_{ivp1} = \begin{cases} 1 & \text{if } Knowledge^*_{ivp1} \leq \nu_1; \\ 2 & \text{if } \nu_1 < Knowledge^*_{ivp1} \leq \nu_2; \\ 3 & \text{if } \nu_2 < Knowledge^*_{ivp1}. \end{cases} \tag{3}$$

$Knowledge_{ivp1} \in \{1,2,3\}$ represents the subject's reported level of RTR understanding at endline, while $Knowledge^*_{ivp1}$ is the corresponding latent variable. Other denotations have the same meaning as in equation (1), and the subject's initial level of RTR understanding reported at baseline is incorporated in the vector of individual covariates X_{ivp0}. The sign of β indicates whether assignment to a given treatment arm makes subjects more or less likely to report a higher understanding of RTR. We also estimate marginal effects for each treatment arm on the likelihood of choosing the most complete description of RTR's meaning (i.e., $Knowledge_{ivp1} = 3$).

Willingness to participate (short-run, self-reported)

We estimate equation (1) where the outcome is self-reported willingness to participate, measured by a Likert scale on a scale of 0 to 10. Self-reported willingness to participate is ordinal in essence. Therefore, we also estimated an ordered logit model for this dependent variable,

see Table A7.

Enrolment (long-run, behavioural outcome)

To measure longer-run behavioural impacts, we analyse an enrolment indicator collected roughly ten months after the intervention. We estimate ITT effects using the same empirical framework as in equation (1), replacing y_{ivp1} with the long-run enrolment outcome. Given the binary nature of enrolment, we use OLS as a linear probability model, so coefficients can be interpreted as percentage-point changes in enrolment relative to the omitted control group; standard errors remain clustered at the village level.

6 Results

This section reports our experimental findings in three steps. We begin with balance checks to validate randomisation. We then verify the 'first stage' of our intervention by examining impacts on RTR-related knowledge (programme literacy). Next, we analyse the divergence between stated and revealed preferences: we first study short-run self-reported willingness to participate (measured immediately post-intervention) and then examine longer-run verified enrolment (measured ten months later).

A key feature of our design is a 3×2 factorial structure: three information-salience contents (TE/BE/HA) crossed with two messenger conditions (Messenger vs. No Messenger). To improve statistical power and transparently map estimates to our research questions, we report pooled ITT estimates along each dimension. In 'Panel A' (Information Salience), we pool across messenger conditions to identify the average effect of content. Symmetrically, in 'Panel B' (Messenger effect), we pool across content categories to isolate the effect of addressing trust friction.

6.1 Balance check

Table 2 presents the balance checks. Panels A and B report baseline RTR-related variables and demographics, respectively. Columns (1) and (2) show control and treatment means,

while Columns (3) and (4) report differences and normalised differences. Column (5) provides Romano–Wolf step-down p-values to adjust for multiple hypothesis testing.

Overall, the randomisation was successful. The treatment and control groups are well-balanced along observed dimensions. While a few unadjusted differences appear statistically significant, the corresponding normalised differences are uniformly small and well below the 0.25 rule-of-thumb for economically meaningful imbalance (Imbens & Rubin, 2015).[5] Romano–Wolf adjusted p-values are large ($p > 0.1$), and an omnibus test fails to reject joint equality of baseline covariates. To further improve precision, we include baseline demographics, RTR experiences, and baseline outcome measures as covariates in all subsequent regressions.

Table 2: Balance check

Variable	Mean Control	Mean Treated	Diff.	Norm. Diff.[1]	Romano-Wolf p-value[2]
Panel A: RTR related results					
Self-reported likelihood of RTR participation	8.403	8.370	-0.033	-0.012	0.999
	(1.756)	(1.972)	(0.839)		
Participation in a toilet retrofitting programme	0.500	0.446	-0.054	-0.077	0.948
	(0.501)	(0.497)	(0.197)		
Participation in a Septic construction programme	0.414	0.419	0.005	0.008	0.999
	(0.494)	(0.494)	(0.901)		
Heard RTR from village committee	0.559	0.576	0.017	0.024	0.995
	(0.498)	(0.495)	(0.598)		
Most crucial determinant for participation					
Cost	0.339	0.364	0.025	0.038	0.994
	(0.475)	(0.482)	(0.521)		
Benefit	0.156	0.129	-0.027	-0.055	0.994
	(0.364)	(0.335)	(0.365)		
Administrative procedure	0.075	0.116	0.041	0.099	0.830
	(0.265)	(0.321)	(0.101)		
Techniques	0.086	0.118	0.032	0.074	0.945
	(0.281)	(0.322)	(0.181)		
Purpose	0.344	0.273	-0.071*	-0.109	0.822

Continued on next page...

[5] See also A. Banerjee, Ferrara, and Orozco (2018).

Table 2: Balance check (Continued)

Variable	Mean Control	Mean Treated	Diff.	Norm. Diff.[1]	Romano-Wolf p-value[2]
	(0.476)	(0.446)	(0.093)		
Understanding of RTR					
Shallow understanding	0.296	0.390	0.095*	0.142	0.730
	(0.458)	(0.488)	(0.078)		
Medium understanding	0.226	0.185	-0.041	-0.071	0.958
	(0.419)	(0.389)	(0.355)		
Good understanding	0.478	0.424	-0.054	-0.077	0.994
	(0.501)	(0.495)	(0.316)		
Level of reliance on government					
No reliance	0.511	0.559	0.049	0.069	0.989
	(0.501)	(0.497)	(0.328)		
Low reliance	0.446	0.343	-0.103**	-0.150	0.547
	(0.498)	(0.475)	(0.034)		
Heavy reliance	0.043	0.098	0.055**	0.152	0.400
	(0.203)	(0.297)	(0.019)		
Panel B: Demographics					
Gender dummy (Female = 1)	0.290	0.285	-0.005	-0.008	0.999
	(0.455)	(0.452)	(0.873)		
Age group					
Adult (18-40)	0.075	0.079	0.004	0.009	0.999
	(0.265)	(0.270)	(0.888)		
Middle age adult (41-61)	0.677	0.641	-0.037	-0.055	0.994
	(0.469)	(0.480)	(0.384)		
Senior adult (61+)	0.247	0.280	0.033	0.053	0.994
	(0.433)	(0.449)	(0.365)		
High school degree above	0.070	0.110	0.040*	0.100	0.830
	(0.256)	(0.313)	(0.098)		
Annual household income greater than 50,000	0.769	0.663	-0.105**	-0.166	0.479
	(0.423)	(0.473)	(0.026)		
Household size greater than 3	0.360	0.395	0.035	0.051	0.994
	(0.481)	(0.489)	(0.435)		
Mandarin speaker	0.462	0.443	-0.019	-0.027	0.995
	(0.500)	(0.497)	(0.691)		
College student living in the house	0.204	0.229	0.025	0.042	0.994
	(0.404)	(0.420)	(0.490)		
Household head	0.602	0.623	0.021	0.031	0.995

Continued on next page...

Table 2: Balance check (Continued)

Variable	Mean Control	Mean Treated	Diff.	Norm. Diff.[1]	Romano-Wolf p-value[2]
	(0.491)	(0.485)	(0.635)		

Notes: Standard deviations are reported in parentheses for the means. Standard errors for the differences in Column (3) are clustered at the village level (unit of randomisation).

[1] Std Diff $= \frac{\bar{X}_{\text{treatment}} - \bar{X}_{\text{control}}}{\text{SE}_{\text{cluster}}}$.

[2] Romano-Wolf p-value after adjusting for multiple hypothesis testing with 1000 repetitions.

6.2 Effects on programme literacy

We first examine if the subjects successfully caught the conveyed information in the treatment videos. Table 3 reports these results with two complementary measures for participants' understanding of the RTR programme. Column (1) uses a dynamic measure: an indicator for Improved Understanding, equal to one if a respondent reports a higher-order category of understanding at endline relative to baseline. Column (2) uses the Level of Understanding at endline (ordinal scale), estimated via an ordered logit model controlling for baseline knowledge.

Panel A confirms that the informational nudges were effective learning aids. Emphasising benefits (BE) or harms (HA) significantly increases knowledge acquisition. In Column (1), the BE and HA arms raise the probability of improving understanding by 6.6 and 11.0 percentage points, respectively, relative to the active placebo. In Column (2), these same arms shift the distribution of endline understanding upward (ordered-logit coefficients of 0.690^{**} and 0.936^{***}). Panel B (messenger) indicates that both messenger presence and no messenger presence increase (1) the probability of within-person improvement (by 8.7 and 6.3 percentage points in Column 1) and (2) the ordered measure of endline understanding (coefficients 0.784 and 0.542 in Column 2), with somewhat larger point estimates when a messenger is present. Marginal effects suggest that these interventions increase the probability of selecting the most comprehensive category by roughly 2–5 percentage points (see Table A3). These 'first-stage' results show that the treatments did successfully shift cognitive understanding.

We further assess whether treated subjects attended to the highlighted attribute using two

Table 3: Treatment effects on knowledge

	(1) Improved Understanding	(2) Level of RTR Understanding
Panel A: Information Salience Treatments		
Technical attributes	0.052*	0.376
	(0.027)	(0.273)
Benefit attributes	0.066**	0.690**
	(0.030)	(0.274)
Harm attributes	0.110***	0.936***
	(0.035)	(0.312)
Observations	985	985
R-squared	0.205	0.588
Individual Control	YES	YES
Village Control	YES	YES
Town FE	YES	YES
Panel B: Messenger Effect Treatments		
Messenger presence	0.087***	0.784***
	(0.023)	(0.206)
No messenger presence	0.063**	0.542**
	(0.025)	(0.243)
Observations	985	985
R-squared	0.201	0.589
Individual Control	YES	YES
Village Control	YES	YES
Town FE	YES	YES

Note: Standard errors are in parentheses, clustered at the village level (unit of randomisation). Column (1) uses an individual-level pre–post change indicator for RTR understanding (improvement relative to baseline). Column (2) reports ordered logit regression results for endline understanding levels (conditional on baseline understanding and baseline information demand included in Individual Control). In Column (2), the Pseudo R^2 is reported. *** p<0.01, ** p<0.05, * p<0.1.

attention-check questions.[6] Among non-attriters in the treatment arms, around 6 per cent failed to identify the main content of the treatment videos. The proportions were 3.98 per cent in TE, 6.67 per cent in BE, and 7.83 per cent in HA, with no statistically significant differences across information-salience arms (joint χ^2 test $p = 0.242$). Comparable results emerge in the full sample.

6.3 Effects on RTR engagement

We next turn to our primary outcomes, distinguishing between short-run stated intentions and longer-run realised behaviour.

[6]Only treated subjects are asked to complete the attention-check questions. Question 1: What does this video talk about? Answers: 1) The benefits of the toilet revolution; 2) The harms caused by unsanitary latrine pits and poor living environment; 3) Technical details for the construction, use and maintenance of sanitary toilets and harmless septic. Question 2: Who is the presenter calling for engagement in RTR that appeared at the beginning of the video? Answers: 1) No one appears; 2) A role model villager who participated in the toilet revolution.

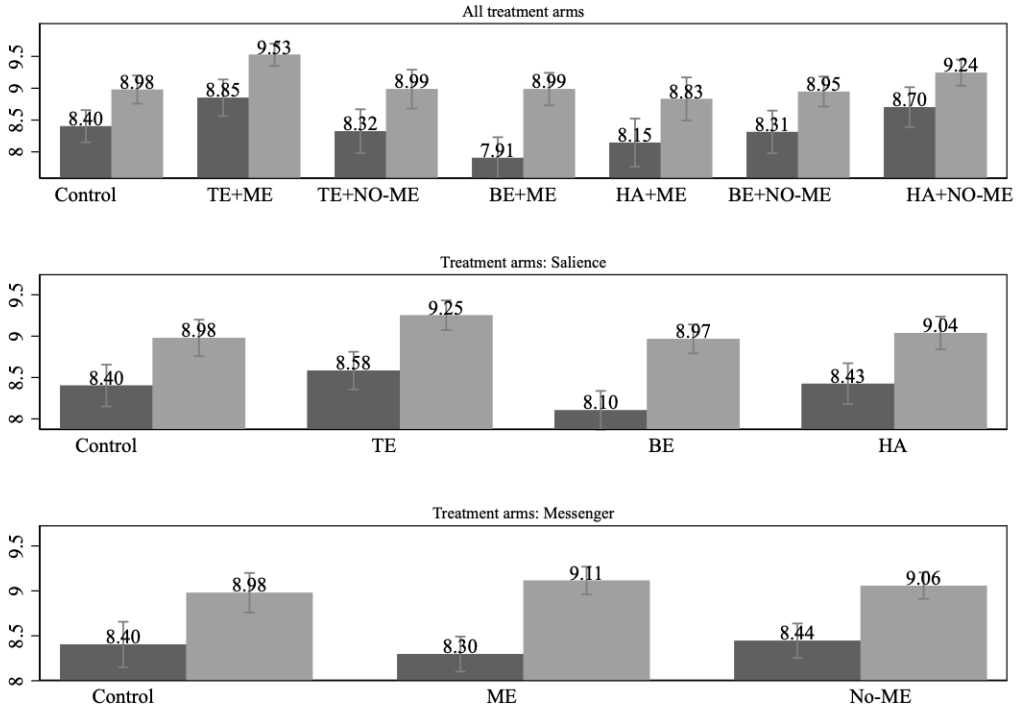

Figure 3: Within group change in willingness to participate

Notes: This figure displays the average self-reported attitude towards RTR programmes (on a scale of 1-10) for the Control group and various treatment combinations, before and after the intervention (the dark grey bars represent the pre-experiment average and the light grey bars represent the post-experiment average). The top panel shows effects for combinations of information salience and messenger effects treatments. The bottom panel shows the average effect of Messenger (ME: Messenger presence, No-ME: No messenger presence). Error bars represent 95% confidence intervals. All values represent the mean attitude score reported by respondents in each group.

Short-run stated willingness: descriptive pre–post patterns We begin with descriptive patterns of willingness to participate measured immediately before and after the video interventions. As shown in Figure 3, average willingness increases substantially over the survey window across all groups, including the active-placebo control group. Within every arm, paired t-tests and Wilcoxon signed-rank tests reject the equality of pre- and post-intervention distributions ($p < 0.001$). However, these broad increases are difficult to interpret as treatment-specific persuasion. The fact that the control group, who viewed a generic video about rural revitalisation without specific sanitation details, experienced a similar boost suggests a strong priming or generic attention. Mere engagement with a similar topic (to RTR) appears sufficient to elevate short-run stated intentions.

Short-run stated willingness: ITT We next estimate equation (1) with endline willingness as the dependent variable and baseline willingness as a covariate.

Table 4 presents the ITT estimates for endline stated willingness. Consistent with the descriptive patterns, we verify a null result in differential impact. Across all specifications in Panel A (Information Salience) and Panel B (Messenger), point estimates are positive but statistically indistinguishable from the active-placebo control. Given the high endline mean in the control group (8.98/10), these results point to a ceiling effect where generic salience maximises stated intent, masking any specific impact of resolving cognitive barriers.[7]

[7]Results remain largely consistent when restricting the sample to non-attriters (see Table A8). While the technical attributes (TE) treatment shows a marginal increase (0.245) in this subsample, it is significant only at the 10% level.

Table 4: Short-term effects on willingness to participate

	(1)	(2)	(3)	(4)
Panel A: Information Salience Treatments				
Technical attributes	0.191	0.200	0.239*	0.220
	(0.128)	(0.127)	(0.123)	(0.140)
Benefit attributes	0.126	0.088	0.096	0.064
	(0.118)	(0.112)	(0.115)	(0.121)
Harm attributes	0.050	0.045	0.039	0.014
	(0.141)	(0.145)	(0.157)	(0.149)
Willingness to participate at baseline	0.460***	0.454***	0.457***	0.462***
	(0.059)	(0.059)	(0.059)	(0.059)
Observations	985	985	985	985
R-squared	0.338	0.361	0.367	0.374
Individual Control	NO	YES	YES	YES
Village Control	NO	NO	YES	YES
Town FE	NO	NO	NO	YES
Panel B: Messenger Effect Treatments				
Messenger presence	0.185	0.161	0.183	0.147
	(0.126)	(0.121)	(0.121)	(0.113)
No messenger presence	0.060	0.063	0.080	0.055
	(0.101)	(0.100)	(0.101)	(0.118)
Willingness to participate at baseline	0.462***	0.456***	0.459***	0.464***
	(0.058)	(0.058)	(0.058)	(0.058)
Observations	985	985	985	985
R-squared	0.338	0.364	0.368	0.373
Individual Control	NO	YES	YES	YES
Village Control	NO	NO	YES	YES
Town FE	NO	NO	NO	YES
Mean Dep. Var. (Control)		8.978		
Std Dev. Dep. Var. (Control)		1.535		

Notes: Standard errors clustered at the village level (unit of randomisation) are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Long-run enrolment We now examine the effects on the verified enrolment roughly ten months later. Unlike the survey measure, this outcome requires real-world effort and is realised outside the immediate framing of the experiment. Table 5 reports the results. In contrast to the short-run null results, we find economically meaningful effects on actual behaviour. Panel A

shows that Benefit Salience (BE) significantly increases long-run enrolment by approximately 9 percentage points (Column 4). This effect is robust to multiple hypothesis testing adjustments (see Table A4). The Harms (HA) arm shows a positive but smaller and statistically insignificant estimate.

It is worth distinguishing the asymmetry between the significant effect of gain-framed messages (Benefit attributes) and the null result of loss-framed messages (Harm attributes). This pattern provides direct evidence regarding our core hypotheses. The significant pooled effect of benefit-related information supports Hypothesis 1 (Salience Effect), confirming that resolving the cognitive valuation gap regarding invisible returns is sufficient to boost take-up. However, regarding the structure of this information, the robust positive effect of gain-framed messages, contrasted with the null effect of loss-framed appeals, leads us to reject Hypothesis 2a (Standard Loss Aversion) in favour of Hypothesis 2b (Hedonic Adaptation). While Prospect Theory suggests that losses often loom larger than gains, our results show that emphasising the 'harms' of the status quo failed to drive enrolment. We attribute this to hedonic adaptation: rural households may have normalised the negative externalities of unimproved sanitation (e.g., odour, flies) as their reference point. Consequently, messages emphasising these conditions are not processed as a 'loss' from a neutral baseline, but rather as a description of the status quo. This finding suggests that policymakers cannot simply assume that the target population perceives the 'badness' of traditional practices as a loss. Interventions attempting to leverage loss aversion may fail if the target audience has normalised the adverse conditions that policymakers seek to eliminate.[8]

Panel B of Table 5 and Table A6 consistently show a null result for the Messenger treatment. Consequently, we find no empirical support for Hypothesis 3 (Messenger Effect). The presence of a peer advocate does not significantly boost enrolment, nor does it moderate the effect of information salience. This result implies that in this context, the binding constraint

[8]This interpretation is corroborated by qualitative evidence from our face-to-face interviews ($N = 14$). Interviews reveal that the negative externalities of traditional latrines are often normalised by residents. For instance, one respondent explicitly noted that he only realised the detrimental health impacts of dirty toilets after watching a specific documentary, implying that prior to this external informational shock, the unsanitary conditions were not perceived as a 'loss' but as a standard living condition. Another informant characterised the pursuit of a hygienic environment as a 'lofty life' secondary to basic livelihood (making money), further suggesting that the current sanitary baseline is accepted as the reference point rather than a deficit to be corrected.

appears to be primarily cognitive (calculation friction regarding utility) rather than normative (social proof). Once the benefits are made salient to the household, the cognitive barrier of benefit calculation is removed, and the additional social cue provides limited marginal value. Alternatively, even if trust friction does exist, our messenger manipulation may have been too low-intensity and insufficiently 'local' to activate peer influence: the messenger's appearance was brief and static, and the messenger was drawn from an adjacent township rather than the recipient villages.[9] This highlights that while standardised mass media can close informational gaps, shifting entrenched practices may require higher-frequency and higher-fidelity social interaction that leverages genuine local trust.

The divergence between stated and revealed preference is the one of the core findings of our study. In a setting where financial costs are zero, the specific resolution of the valuation problem via Benefit Salience was the only lever that translated into action. While generic attention was enough to boost stated willingness, only the clarity provided by the Benefit treatment was sufficient to drive households to navigate the administrative process of enrolment.[10]

[9] In the design stage, we initially considered having the selected villager record a video for the entire introductory part. However, due to educational limitations, the final output lacked fluency and naturalness. Therefore, we ultimately decided to use a static image, which limited the intervention's intensity. The critical point is that it was hard to find someone that everyone recognised in our context, and we acknowledge this decision is one limitation.

[10] To increase precision, we pool the conceptually similar Benefit and Harm arms ('Benefit-Salience') in Table A6. We find that communicating the welfare benefits of sanitation (Gain or Loss framed) increases enrolment by 6.8 percentage points ($p < 0.1$). In addition, we estimate an ordered logit and a logit regression for WTP and enrolment indicator to match the outcome variable type. The results are shown in Table A7 and we obtain consistent conclusions.

Table 5: Long-term effects on enrolment decision

	(1)	(2)	(3)	(4)
Panel A: Information Salience Treatments				
Technical attributes	-0.006	0.011	0.018	0.009
	(0.038)	(0.039)	(0.039)	(0.040)
Benefit attributes	0.099**	0.091**	0.095**	0.088**
	(0.040)	(0.044)	(0.044)	(0.044)
Harm attributes	0.054	0.045	0.045	0.043
	(0.038)	(0.040)	(0.039)	(0.040)
Participated in toilet retrofitting programme at baseline	-0.085**	-0.081*	-0.081*	-0.085*
	(0.040)	(0.043)	(0.043)	(0.043)
Participated in septic construction programme	-0.002	0.003	0.004	0.008
	(0.040)	(0.044)	(0.044)	(0.044)
Observations	730	730	730	730
R-squared	0.015	0.040	0.043	0.046
Individual Control	NO	YES	YES	YES
Village Control	NO	NO	YES	YES
Town FE	NO	NO	NO	YES
Panel B: Messenger Effect Treatments				
Messenger presence	0.046	0.041	0.046	0.042
	(0.036)	(0.038)	(0.037)	(0.039)
No messenger presence	0.045	0.052	0.054	0.051
	(0.035)	(0.035)	(0.035)	(0.036)
Participated in toilet retrofitting programme at baseline	-0.086**	-0.081*	-0.079*	-0.084*
	(0.040)	(0.042)	(0.043)	(0.043)
Participated in septic construction programme	0.004	0.005	0.006	0.010
	(0.040)	(0.043)	(0.043)	(0.043)
Observations	730	730	730	730
R-squared	0.009	0.036	0.040	0.043
Individual Control	NO	YES	YES	YES
Village Control	NO	NO	YES	YES
Town FE	NO	NO	NO	YES
Mean Dep. Var. (Control)		0.331		
Std Dev. Dep. Var. (Control)		0.472		

Notes: Standard errors clustered at the village level (unit of randomisation) are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

6.4 Statistical Power

In this section, we assess the statistical power of our design to detect the actual treatment effect magnitudes observed in the study. We rely on the realised precision from our pooled specifications (results reported in Table A6), which yielded a standard error of approximately 0.036 for benefit-salience interventions regardless of framing. Using the standard parameters ($\alpha = 0.05$, two-sided), we calculate the ex-post power to detect the primary treatment effect of 8.8 percentage points (as observed in the Benefit Salience arm) as follows:

$$Power = \Phi\left(\frac{\delta}{SE} - z_{1-\alpha/2}\right) = \Phi\left(\frac{0.088}{0.036} - 1.96\right) \approx 69\% \tag{4}$$

where Φ is the cumulative distribution function of the standard normal distribution. We calculate a statistical power of approximately 69 percent. Although this is below the standard 80 per cent benchmark, the fact that we detected a significant result confirms that the behavioural shift was distinct enough to be measured by our experimental setup.

6.5 Heterogeneity: The role of baseline programme literacy

In this section, we examine the mechanisms driving the enrolment effects documented in Section 6 by investigating the heterogeneity of treatment effects. Specifically, we ask whether the informational nudge addresses cognitive barriers on the extensive margin or the intensive margin. If the cognitive barrier of valuation problem is primarily driven by the unawareness of the programme's details, the video should be most effective for the least informed households. For them, the benefit-salient information fills a gap and acts as a substitute for their lack of prior knowledge (Extensive margin). Conversely, if the main problem is that the programme's value is too complex to process, then a short video might not be enough for someone starting from zero. Instead, the intervention may only work for those who already have a basic foundation. For these households, the video acts as a complement, helping them connect the dots and tipping their decision toward enrolment (Intensive margin).

To test these hypotheses, we re-estimate the enrolment effects separately for subgroups defined by their baseline RTR understanding. Figure 4 presents the results. We find evidence

of heterogeneity: the Benefit Salience (BE) intervention significantly increases enrolment only among subjects who possessed a foundational level of RTR knowledge prior to the experiment.[11] For this subgroup, the treatment effect is large and statistically significant, whereas for those with low baseline knowledge, the effect is statistically indistinguishable from zero. This pattern strongly supports the intensive margin hypothesis. It also suggests that a brief video highlighting benefits may be insufficient to bridge the gap for those starting from zero. In contrast, for those who already understood the basic concept of the RTR, the intervention successfully resolved the remaining uncertainty regarding the value of the upgrade (the invisible benefits), thereby tipping their decision calculus toward enrolment.

This finding implies that while the interventions successfully improved short-run knowledge scores across the board (as shown in Table 3), this knowledge gain only translated into behavioural change for those with a pre-existing cognitive framework. The positive correlation between baseline and endline knowledge ($p < 0.01$) further reinforces this dynamic: initial knowledge facilitates the acquisition and retention of new, policy-relevant information. Consequently, the mechanism driving our main result appears to be the resolution of a valuation problem among the moderately informed, rather than a simple alerting effect for the uninformed. This has crucial policy implications: it implies that low-cost salience nudges are not a standalone solution for the entire population. To reach the households with low baseline programme literacy, policymakers may need to invest in foundational education campaigns to build the necessary programme literacy before salience interventions can trigger behavioural change.

[11] Subjects with 'foundational RTR understanding' are defined as those who did not select the simplest and most superficial definition of the RTR programme at baseline. Approximately 47% of the sample falls into this category.

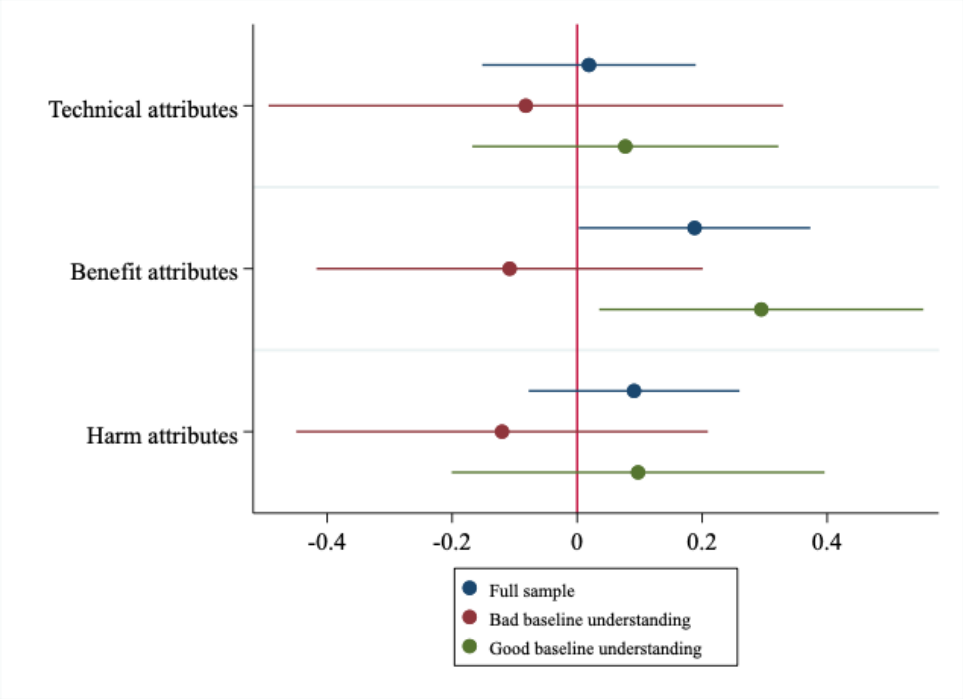

Figure 4: Effect size of salience treatments

Notes: This figure presents the estimated causal effects of the three information salience treatments (Technical, Benefit, and Harm attributes) on probability of engagement in RTR, measured in standard deviations for ease of comparison, reported in Table A5. We disaggregate the full sample to examine heterogeneous treatment effects across two distinct subgroups: those with low and high baseline understanding of the RTR programme. We anticipate that this disaggregation will reveal differences in responsiveness, consistent with models of information demand where marginal returns to information are highest for those with moderate pre-existing knowledge

6.6 Attrition

Our study begins with 985 participants across 93 villages in six towns selected for low coverage of sanitary toilets. The overall attrition rate is 25.89 per cent, with 24.91 per cent in treated villages and 30.11 per cent in control villages. A joint test yields a p-value of 0.145, implying that we cannot reject equality of attrition rates between treatment and control groups at con-

ventional levels.[12] To assess whether attrition poses a threat to internal validity, we proceed in two steps. First, we compare baseline characteristics between attriters and non-attriters. Table A9 reports correlates of attrition and indicates that subjects who participated in previous RTR projects are more likely to drop out. Second, we test for differential attrition across experimental arms. Table A10 reports adjusted p-values from balance tests comparing attriters in each treatment arm to attriters in the control group. We find no evidence of differential attrition across arms, suggesting that treatment–control comparisons within the observed (non-attriting) sample remain informative for our main conclusions.

7 Discussion

Our findings provide new insights to explain the persistence of incomplete take-up for welfare-improving programmes even when monetary costs are negligible. Re-validating the puzzle highlighted at the outset, we show that removing financial barriers is insufficient to guarantee participation. By exploiting the unique "zero-cost" laboratory of China's Rural Toilet Revolution, we isolate non-monetary frictions and confirm that the binding constraint is primarily a cognitive valuation problem: rural households appear to face cognitive barriers that impede their ability to quantify the invisible returns of adoption. Our results show that breaking this barrier requires interventions that specifically make these invisible benefits salient, thereby closing the valuation gap. This finding provides robust empirical support for Hypothesis 1 (Salience Effect), confirming that in the absence of financial constraints, information is not merely facilitating, it is foundational.

A central finding of this study is the competing behavioural mechanisms regarding information framing. Standard Prospect Theory predicts that highlighting the harms of the status quo should leverage loss aversion to drive adoption (H2a). However, our results firmly reject this standard prediction in favour of Hypothesis 2b (Hedonic Adaptation). While the gain-framed intervention significantly increased verified enrolment, the loss-framed interven-

[12] While the attrition rate may appear high, our collaboration protocol with the local authority prohibits us from collecting personal identifiers other than villagers' telephone numbers. This limitation reduces our ability to track subjects over time. We nonetheless implemented multiple mitigation strategies (e.g., administrative assistance from village committees and manual matching of contact numbers) to minimise attrition.

tion yielded null results. We attribute this asymmetry to the shifting of reference points under chronic deprivation. For households habituated to traditional latrines, the negative externalities (e.g., filth, disease) are likely processed not as a "loss" from a healthy norm, but as the neutral status quo (Frederick & Loewenstein, 1999). Consequently, loss-framed appeals failed to trigger the expected aversion because the audience perceived no deviation from their baseline. In contrast, gain-framed messages succeeded by projecting a modernised lifestyle, effectively creating an aspirational gap above the adapted baseline. This finding offers a critical refinement to Behavioural Public Policy: in contexts of normalised deprivation, "scare tactics" may be cognitively inert; effective nudges must instead construct aspirational reference points.

Furthermore, regarding social proximity, we find no empirical support for Hypothesis 3 (Messenger Effect). The presence of a peer advocate did not significantly boost enrolment in our context, suggesting that the resonance gap was either not the binding constraint or was not successfully bridged by our intervention. This null result implies a hierarchy of frictions: in this context, the primary barrier is cognitive (calculation friction regarding utility) rather than normative (social proof). Once the benefits were made salient (resolving the cognitive barrier), the additional social cue provided limited marginal utility. Alternatively, as noted in the limitations, the low intensity of a video-based messenger may have been insufficient to generate the "social proximity" required to alter behaviour, highlighting the distinction between mass-media scalability and high-fidelity social influence.

Our study also underscores a critical methodological caution regarding the reliance on stated preferences. We observe a sharp divergence between the uniform increase in short-run self-reported willingness and the specific treatment effects on long-run verified enrolment (costly action). This discrepancy suggests that superficial survey responses may mask the true binding constraints. While generic advocacy can elicit performative support via priming or experimenter demand effects, only interventions that specifically resolve the cognitive valuation gap translate into real-world behavioural change.

From a policy perspective, the heterogeneous treatment effect by baseline programme literacy points to a complementarity between information and baseline programme knowledge. We find that salience nudges operate on the intensive margin: they are highly effective for house-

holds with a foundational understanding of the programme but fail to move the uninformed group. This implies a stage-dependent policy strategy. For the least literate populations, low-cost nudges are insufficient. However, once a basic understanding is established, informational nudges become highly cost-effective tools to bridge the last mile of adoption.

Two features of our study design are important for interpreting external validity. First, the study was implemented in a region where the RTR has been actively rolled out but coverage is incomplete. Therefore, our estimates are most naturally viewed as applying to the expansion phase of public programmes, where the low-hanging fruit (early adopters) may have been harvested, but a large "persuadable" middle group remains. Second, the zero-cost institutional setting is specific but not unique; it mirrors many turnkey public infrastructure projects in developing countries (e.g., last-mile electrification or water connection). In such settings, our results imply that demand-side resources should focus less on subsidies (which are already maximised) and more on bridging the cognitive gap between infrastructure and household utility.

Finally, we perform a back-of-the-envelope calculation to benchmark the economic viability of digital advocacy. We parameterise the one-time fixed cost (C) of producing high-quality advocacy videos between 10,000 and 90,000 RMB (approx. \$1,400–\$12,300) based on market rates for animation and filming. On the benefit side, we rely on Wang and Shen (2022), who estimate that the Rural Toilet Revolution increases household labour supply by 0.25–0.43 hours per day; valued at the local minimum wage, this translates to an annual income gain (g) of approximately 1,214–2,069 RMB per participating worker.

Applying our pooled enrolment effect of 6.8 percentage points to the study sample ($N = 985$), the intervention generates approximately 67 induced participants, yielding an aggregate annual income gain of 81,365 to 138,609 RMB for the treated villages. Comparing these benefits against the fixed costs reveals a highly favourable payback horizon. Under central estimates (average production cost of 50,000 RMB and average wage gains), the economic returns generated by new enrollees cover the production costs in approximately 0.45 years (5.4 months). Even under a conservative 'worst-case' scenario (highest costs, lowest wage returns, and lower-bound enrolment), the investment is recovered within 2.5 years.

Crucially, this calculation applies to a pilot scale of fewer than 1,000 households. Because

the marginal cost of digital dissemination via existing social networks is negligible, the cost-effectiveness improves strictly linearly with scale. Taken together, our results suggest that for turnkey infrastructure projects where liquidity is not binding, shifting demand-side resources from hardware subsidies to targeted, benefit-salient communication is a welfare-enhancing strategy with immense potential for returns to scale.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used GPT-5.2-Thinking (Dec 2025 version) and Claude-4.5-Sonnet (Nov 2025 version) in order to improve the readability and language of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Disclosure statement

The authors report there are no competing interests to declare.

Funding

This work was supported by the National Natural Science Foundation of China (No. 72203054), the Basic and Applied Basic Research Foundation of Guangdong Province (No. 2025A1515010549), the Guangzhou Municipal Science and Technology Project (No. 2024A03J0630 and No. 2024A04J4549), the Yangcheng Scholars Research Project of Guangzhou (No. 2024312031), and the Guangzhou-HKUST(GZ) Joint Funding Program (No. 2025A03J3807).

Data availability statement

The data that support the findings of this study are available upon request.

References

Alsan, M., Garrick, O., & Graziani, G. (2019). Does diversity matter for health? experimental evidence from oakland. *American Economic Review*, *109*(12), 4071–4111.

Alsan, M., & Wanamaker, M. (2018). Tuskegee and the health of black men. *The quarterly journal of economics*, *133*(1), 407–455.

Bai, C.-E., Chi, W., Liu, T. X., Tang, C., & Xu, J. (2021). Boosting pension enrollment and household consumption by example: A field experiment on information provision [Journal Article]. *Journal of Development Economics*, *150*, 102622. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304387820301978 doi: https://doi.org/10.1016/j.jdeveco.2020.102622

Banerjee, A., Ferrara, E. L., & Orozco, V. (2018). Entertainment, education, and attitudes toward domestic violence. In *Aea papers and proceedings* (Vol. 109, p. 133-37).

Banerjee, A. V., Banerji, R., Duflo, E., Glennerster, R., & Khemani, S. (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in india. *American Economic Journal: Economic Policy*, *2*(1), 1–30.

BenYishay, A., & Mobarak, A. M. (2014). *Social learning and communication* (Tech. Rep.). National Bureau of Economic Research.

Bhargava, S., & Manoli, D. (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *American Economic Review*, *105*(11), 3489–3529.

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of personality and social psychology*, *36*(8), 917.

Brookings Institution. (2022). *What's next for poverty reduction policies in china?* https://www.brookings.edu/articles/whats-next-for-poverty-reduction-policies-in-china/. (Accessed: 2025)

Cheng, S., Li, Z., Uddin, S. M. N., Mang, H.-P., Zhou, X., Zhang, J., ... Zhang, L. (2018). Toilet revolution in china [Journal Article]. *Journal of Environmental Management*, *216*, 347-356. Retrieved from https://pubmed.ncbi.nlm.nih.gov/28941832https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5937855/ doi: 10.1016/j.jenvman.2017

.09.043

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American economic review*, *99*(4), 1145–1177.

Cohen, J., & Dupas, P. (2010). Free distribution or cost-sharing? evidence from a randomized malaria prevention experiment. *The Quarterly Journal of Economics*, *125*(1), 1–45.

Currie, J. (2004). *The take up of social benefits.* National Bureau of Economic Research Cambridge, Mass., USA.

Dupas, P. (2011). Health behavior in developing countries. *Annu. Rev. Econ.*, *3*(1), 425–449.

Frederick, S., & Loewenstein, G. (1999). Hedonic adaptation.

Ganzach, Y., & Karsahi, N. (1995). Message framing and buying behavior: A field experiment. *Journal of Business Research*, *32*(1), 11-17. Retrieved from `https://www.sciencedirect.com/science/article/pii/0148296393000383` doi: https://doi.org/10.1016/0148-2963(93)00038-3

Geruso, M., & Spears, D. (2018, April). Neighborhood sanitation and infant mortality. *American Economic Journal: Applied Economics*, *10*(2), 125-62. Retrieved from `https://www.aeaweb.org/articles?id=10.1257/app.20150431` doi: 10.1257/app.20150431

Guiteras, R., Levinsohn, J., & Mobarak, A. M. (2015). Encouraging sanitation investment in the developing world: A cluster-randomized trial. *Science*, *348*(6237), 903–906.

Handel, B., & Schwartzstein, J. (2018). Frictions or mental gaps: what's behind the information we (don't) use and when do we care? *Journal of Economic Perspectives*, *32*(1), 155–178.

Homar, A. R., & Cvelbar, L. K. (2021). The effects of framing on environmental decisions: A systematic literature review. *Ecological Economics*, *183*, 106950.

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, *125*(2), 515–548.

Kahneman, D., Tversky, A., et al. (1979). Prospect theory: An analysis of decision under risk.

Econometrica, *47*(2), 363–391.

Kar, K., & Pasteur, K. (2005). Subsidy or self-respect?: Community led total sanitation; an update on recent developments.

Ko, W., & Moffitt, R. A. (2022, June). *Take-up of social benefits* (Working Paper No. 30148). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w30148 doi: 10.3386/w30148

Kremer, M., Rao, G., & Schilbach, F. (2019). Behavioral development economics. *Handbook of Behavioral Economics - Foundations and Applications 2*.

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, *76*(2), 149-188. Retrieved from https://www.sciencedirect.com/science/article/pii/S0749597898928047 doi: https://doi.org/10.1006/obhd.1998.2804

Li, Y., Cheng, S., Li, Z., Song, H., Guo, M., Li, Z., ... Li, T. (2021). Using system dynamics to assess the complexity of rural toilet retrofitting: Case study in eastern china [Journal Article]. *Journal of Environmental Management*, *280*, 111655. Retrieved from https://www.sciencedirect.com/science/article/pii/S0301479720315802 doi: https://doi.org/10.1016/j.jenvman.2020.111655

Lichand, G., & Mani, A. (2020). Cognitive droughts. *University of Zurich, Department of Economics, Working Paper*(341).

Liu, Y., Zang, Y., & Yang, Y. (2020). China's rural revitalization and development: Theory, technology and management. *Journal of Geographical Sciences*, *30*, 1923–1942.

Loyalka, P., Liu, C., Song, Y., Yi, H., Huang, X., Wei, J., ... Rozelle, S. (2013). Can information and counseling help students from poor rural areas go to high school? evidence from china [Journal Article]. *Journal of Comparative Economics*, *41*(4), 1012-1025.

Manoli, D. S., & Turner, N. (2014). *Nudges and learning: Evidence from informational interventions for low-income taxpayers* (Report). National Bureau of Economic Research.

Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much*. Macmillan.

Ruggeri, K., Alí, S., Berge, M. L., Bertoldo, G., Bjørndal, L. D., Cortijos-Bernabeu, A., ... others (2020). Replicating patterns of prospect theory for decision under risk. *Nature human behaviour*, *4*(6), 622–633.

Stockman, L. J., Fischer, T. K., Deming, M., Ngwira, B., Bowie, C., Cunliffe, N., ... Quick, R. E. (2007). Point-of-use water treatment and use among mothers in malawi. *Emerging infectious diseases*, *13*(7), 1077.

Wang, D., & Shen, Y. (2022). Sanitation and work time: Evidence from the toilet revolution in rural china. *World Development*, *158*, 105992. Retrieved from https://www.sciencedirect.com/science/article/pii/S0305750X22001826 doi: https://doi.org/10.1016/j.worlddev.2022.105992

World Bank. (2022). *Lifting 800 million people out of poverty – new report looks at lessons from china's experience.* https://www.worldbank.org/en/news/press-release/2022/04/01/lifting-800-million-people-out-of-poverty-new-report-looks-at-lessons-from-china-s-experience. (Accessed: 2025)

A Tables

Table A1: Attribute explanation

Attribute	Explanation
Programme purpose	The intended target to be achieved by the programme, e.g., cleaner living environment and higher life quality for villagers
Application procedure	
Application eligibility	Conditions to be met to be eligible for the programme
Registration channel	On-site registration or online registration, and materials to be submitted
Costs	
Pecuniary costs	Estimated amount of money to be paid by registered households
Time costs	Estimated time consumed to complete the programme
Physical efforts	Estimated physical efforts to be put into the programme
Benefits	
Pecuniary benefits	Amount and form of financial benefits
Other benefits	Non-pecuniary benefits other than financial gains
Improved life quality	Cleaner living environment
Health benefits	Prevention of diseases and public health
Required skills	
Construction skills	Skills required for the construction and renovation of sanitation facilities
Skills to use and manage the facility	Skills required for routine use and maintenance of the sanitation facility

Table A2: Village summary statistics

	N	Mean	SD	Max	Min
Total Population	93	2,453.763	1,259.388	10,380	692
Households	93	880.226	543.224	4,601	225
Households Size ≥ 3	93	0.430	0.498	1	0
Sex Ratio (Female %)	93	0.487	0.020	0.530	0.386
Poor Households (%)	93	0.078	0.049	0.260	0.012
Disposable Income (10,000 Yuan)	93	1.487	0.307	2.800	0.830
Area (km^2)	93	11.719	14.654	72.840	0.150

Note: Descriptive statistics for the 93 villages included in the sample.

Table A3: Marginal effects of interventions on reporting the most complete RTR understanding

	Full sample	**Non-attriter sample**
	(1)	(2)
Panel A: Information salience interventions		
Technical attributes	0.016	0.017
	(0.012)	(0.014)
Benefit attributes	0.029**	0.029**
	(0.013)	(0.015)
Harm attributes	0.040***	0.051***
	(0.014)	(0.017)
Observations	985	730
Panel B: Pooled benefit interventions		
Technical attributes	0.016	0.017
	(0.012)	(0.014)
Pooled benefits	0.034***	0.039***
	(0.010)	(0.013)
Observations	985	730
Panel C: Messenger effects interventions		
Messenger presence	0.033***	0.036***
	(0.010)	(0.012)
No messenger presence	0.023**	0.026*
	(0.011)	(0.014)
Observations	985	730

Notes: Robust standard errors clustered at the village level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Adjusted p-value

	Model p-value	Romano-Wolf p-value
Panel A: Information Salience Treatment		
Dependent Variable: Willingness to Participate		
Technical attributes	0.1196	0.2238
Benefit attributes	0.5994	0.8741
Harm attributes	0.9254	0.9421
Dependent Variable: Enrolment Decision		
Technical attributes	0.8263	0.9421
Benefit attributes	0.0469	0.0839
Harm attributes	0.2863	0.5315
Panel B: Pooled Benefit Treatment		
Dependent Variable: Willingness to Participate		
Technical attributes	0.1193	0.1309
Pooled benefits	0.7169	0.8721
Dependent Variable: Enrolment Decision		
Technical attributes	0.8312	0.8721
Pooled benefits	0.0608	0.0819

Note: Romano-Wolf step-down adjusted p-values are calculated with 1,000 repetitions for re-sampling. Standard errors are clustered at the village level

Table A5: Heterogeneous treatment effects by baseline understanding

	(1) **Full Sample**	(2) **Low Baseline Understanding**	(3) **High Baseline Understanding**
Technical attributes	0.019	-0.082	0.077
	(0.086)	(0.207)	(0.123)
Benefit attributes	0.188**	-0.108	0.295**
	(0.093)	(0.155)	(0.130)
Harm attributes	0.091	-0.120	0.098
	(0.085)	(0.165)	(0.150)
Observations	730	276	454
Baseline Understanding Control	YES	YES	YES
Individual Control	YES	YES	YES
Village Control	YES	YES	YES
Town FE	YES	YES	YES

Notes: Robust standard errors clustered at the village level in parentheses. The dependent variable is verified enrolment. Columns (2) and (3) split the sample based on baseline RTR understanding: 'Low' refers to subjects who chose the most superficial definition, while 'High' refers to those with foundational initial understanding. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A6: Confirmation of main results

	Willingness to participate	**Enrolment decision**
	(1)	(2)
Panel A: Pooled benefits interventions		
Technical attributes	0.220	0.009
	(0.140)	(0.041)
Benefits (Gain & Loss-framed)	0.040	0.068*
	(0.111)	(0.036)
Willingness to participate at baseline	0.461***	-0.010
	(0.059)	(0.010)
Participated in toilet retrofitting programme at baseline	0.102	-0.085*
	(0.091)	(0.043)
Participated in septic construction programme	0.040	0.006
	(0.092)	(0.043)
Observations	985	730
R-squared	0.374	0.045
Individual Control	YES	YES
Village Control	YES	YES
Town FE	YES	YES
Panel B: Moderating role of messenger effects		
Technical attributes	0.109	0.027
	(0.189)	(0.051)
Benefits (Gain & Loss-framed)	0.019	0.068*
	(0.116)	(0.040)
Messenger × Technical attributes	0.211	-0.034
	(0.160)	(0.067)
Messenger × Benefits (Gain & Loss-framed)	0.036	0.001
	(0.149)	(0.045)
Willingness to participate at baseline	0.460***	-0.010
	(0.060)	(0.010)
Participated in toilet retrofitting programme at baseline	0.111	-0.086*
	(0.090)	(0.044)
Participated in septic construction programme	0.029	0.008
	(0.094)	(0.043)
Observations	985	730
R-squared	0.376	0.045
Individual Control	YES	YES
Village Control	YES	YES
Town FE	YES	YES

Table A7: Robustness check: Outcome type

	(1)	(2)
	WTP	**True participation decision 10 months after treatment**
Technical attributes	0.455*	0.045
	(0.267)	(0.183)
Benefit attributes	-0.049	0.394**
	(0.260)	(0.189)
Harm attributes	0.150	0.204
	(0.238)	(0.175)
Observations	985	730
Individual Control	YES	YES
Village Control	YES	YES
Town FE	YES	YES

Notes: Robust standard errors clustered at the village level in parentheses. Column (1) reports ordered logit regression coefficients for WTP. Column (2) reports logit regression coefficients for the binary enrolment decision. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A8: Impacts on willingness to participate: Non-attriter sample

	(1)	(2)	(3)	(4)
Panel A: Information Salience Treatments				
Technical attributes	0.220	0.218	0.268**	0.245*
	(0.137)	(0.132)	(0.123)	(0.144)
Benefit attributes	0.046	-0.016	-0.006	-0.029
	(0.151)	(0.138)	(0.138)	(0.141)
Harm attributes	0.044	0.037	0.044	0.018
	(0.156)	(0.162)	(0.171)	(0.167)
Willingness to participate at baseline	0.448***	0.437***	0.440***	0.444***
	(0.063)	(0.066)	(0.067)	(0.067)
Observations	730	730	730	730
R-squared	0.338	0.373	0.380	0.383
Individual Control	NO	YES	YES	YES
Village Control	NO	NO	YES	YES
Town FE	NO	NO	NO	YES
Panel B: Messenger Effect Treatments				
Messenger presence	0.172	0.142	0.171	0.139
	(0.141)	(0.133)	(0.129)	(0.129)
No messenger presence	0.042	0.034	0.060	0.026
	(0.121)	(0.117)	(0.115)	(0.131)
Willingness to participate at baseline	0.453***	0.441***	0.445***	0.449***
	(0.062)	(0.066)	(0.067)	(0.067)
Observations	730	730	730	730
R-squared	0.337	0.370	0.377	0.380
Individual Control	NO	YES	YES	YES
Village Control	NO	NO	YES	YES
Town FE	NO	NO	NO	YES
Mean Dep. Var. (Control)		8.961		
Std Dev. Dep. Var. (Control)		1.567		

Notes: Robust standard errors clustered at the village level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Determinants of attrition

	Mean Non-attriter	Mean Attriter	Diff.	Norm. Diff.	Romano-Wolf p-value
Self-reported likelihood of RTR participation	9.074	9.043	-0.031	-0.014	1.000
	(1.503)	(1.632)	(0.804)		
Participation in a toilet retrofitting programme	0.389	0.647	0.258***	0.377	0.001
	(0.488)	(0.479)	(0.000)		
Participation in a septic construction programme	0.348	0.624	0.276***	0.405	0.001
	(0.477)	(0.485)	(0.000)		
Heard of RTR from village committee	0.584	0.533	-0.050	-0.072	0.959
	(0.493)	(0.500)	(0.181)		
Cost	0.151	0.129	-0.021	-0.043	0.991
	(0.358)	(0.336)	(0.373)		
Benefit	0.258	0.322	0.064*	0.100	0.835
	(0.438)	(0.468)	(0.093)		
Administrative procedure	0.168	0.192	0.024	0.043	0.991
	(0.375)	(0.395)	(0.373)		
Techniques	0.082	0.059	-0.023	-0.065	0.962
	(0.275)	(0.236)	(0.210)		
Purpose	0.341	0.298	-0.043	-0.065	0.962
	(0.474)	(0.458)	(0.237)		
Shallow understanding of RTR	0.295	0.310	0.015	0.024	1.000
	(0.456)	(0.463)	(0.689)		
Medium understanding of RTR	0.233	0.204	-0.029	-0.050	0.991
	(0.423)	(0.404)	(0.360)		
Good understanding of RTR	0.473	0.486	0.014	0.019	1.000
	(0.500)	(0.501)	(0.714)		
No reliance	0.586	0.600	0.014	0.020	1.000
	(0.493)	(0.491)	(0.731)		
Low reliance	0.377	0.365	-0.012	-0.018	1.000
	(0.485)	(0.482)	(0.765)		
Heavy reliance	0.037	0.035	-0.002	-0.006	1.000
	(0.189)	(0.185)	(0.900)		
Gender dummy (Female = 1)	0.274	0.322	0.048	0.074	0.959
	(0.446)	(0.468)	(0.197)		
Adult	0.078	0.078	0.000	0.001	1.000
	(0.268)	(0.269)	(0.987)		
Middle age adult	0.651	0.639	-0.011	-0.017	1.000
	(0.477)	(0.481)	(0.716)		
Senior adult	0.271	0.282	0.011	0.018	1.000
	(0.445)	(0.451)	(0.688)		
High school degree above	0.096	0.122	0.026	0.058	0.962
	(0.295)	(0.327)	(0.218)		
Annual household income greater than 50,000	0.696	0.651	-0.045	-0.068	0.962
	(0.460)	(0.478)	(0.255)		
Household size greater than 3	0.375	0.427	0.052	0.075	0.917
	(0.485)	(0.496)	(0.136)		
Mandarin speaker	0.425	0.506	0.081**	0.115	0.489
	(0.495)	(0.501)	(0.031)		
Have college student living in the house or not	0.201	0.286	0.085**	0.140	0.243
	(0.401)	(0.453)	(0.011)		
Be the household head or not	0.632	0.584	-0.047	-0.068	0.962
	(0.483)	(0.494)	(0.213)		
Observations	730	255	985		

Notes: Standard deviations are in parentheses under the mean columns. P-values for tests of

Table A10: Romano-Wolf p-values for differential attrition

	Control =T1	**Control =T2**	**Control =T3**	**Control =T4**	**Control =T5**	**Control =T6**
Self-reported likelihood of RTR participation	0.999	0.992	0.856	1.000	1.000	0.999
Participated in a toilet retrofitting programme	0.421	0.909	0.995	1.000	1.000	0.547
Participated in a septic construction programme	0.829	0.653	1.000	1.000	0.997	0.814
Heard RTR from village committee	0.991	0.995	1.000	1.000	0.999	0.999
Cost	0.869	0.907	0.983	1.000	0.999	1.000
Benefit	1.000	0.630	0.995	0.998	0.999	0.982
Administrative procedure	0.784	0.936	0.694	0.704	0.953	0.547
Techniques	0.998	1.000	0.995	1.000	0.999	0.653
Purpose	1.000	1.000	0.981	1.000	0.974	1.000
Shallow understanding of RTR	1.000	0.909	0.989	0.979	0.671	1.000
Medium understanding of RTR	0.991	0.995	0.999	0.967	0.997	1.000
Good understanding of RTR	1.000	0.909	0.795	1.000	0.994	0.999
No reliance	1.000	0.203	0.995	0.974	0.587	1.000
Low reliance	0.850	0.296	0.993	0.977	0.803	1.000
Heavy reliance	0.766	0.614	1.000	1.000	0.999	1.000
Gender (Female = 1)	0.903	1.000	1.000	0.974	0.999	0.986
Adult (18-40)	1.000	1.000	1.000	1.000	1.000	1.000
Middle age adult (41-61)	1.000	1.000	0.995	0.885	1.000	1.000
Senior adult (61 and above)	1.000	0.995	0.995	0.910	1.000	1.000
High school degree above	0.929	0.949	0.999	0.519	0.980	0.821
Annual household income greater than 50,000	0.717	0.211	0.989	0.361	0.173	0.858
Household size greater than 3	0.988	0.998	0.989	1.000	0.953	1.000
Mandarin speaker	1.000	0.995	0.999	1.000	1.000	1.000
Have college student living in the house or not	1.000	0.965	1.000	0.998	0.846	0.982
Household head	0.633	0.995	0.694	0.993	0.994	1.000

Notes: This table reports Romano-Wolf step-down adjusted p-values with 1,000 repetitions. 'Control=Tj' indicates a differential-attrition balance test comparing attriters in the control group to attriters in group Tj

B Video scripts

Note: The scripts below reproduce the content shown to participants.

1. **Introduction to the toilet revolution**

   Since the rural toilet revolution was launched in 2015 in China, the coverage rate of sanitary toilets in rural China has increased significantly: by 2020, it reached over 68%. The rural toilet revolution is not only about the construction and renovation of rural toilets, but also a cleaning movement aimed at modernising rural environmental infrastructure and services, and improving the living environment and life quality for rural residents.

   The rural toilet revolution is essentially a revolution in everyday living habits. It is an important measure to improve the rural living environment and to revitalise rural culture and ecology. We interviewed one participant to share his thoughts, and he emphasised that 'participating in the toilet revolution brings us a healthier and more comfortable living environment. I hope you will join us in the toilet revolution and actively help maintain and improve rural sanitation facilities!' (the underscored sentence will be read by the interviewer in PI treatment groups while read by a voiceover in NON-ME treatments).

2. **Technical attributes (TA)**

   Overview

   The sanitary toilets we are now promoting are generally flush squat toilets, and the faeces can be directly discharged into sewers, biogas tanks, or three-chamber septic tanks. However, some residents still use insanitary pit latrines or directly dig pits to bury their excrement. The sanitary toilets we promote are hygienic, safe, and convenient, and their cleaning and maintenance are also very simple.

   At the same time, we are actively promoting the construction of 'three-chamber harmless septic systems'. 'Three-chamber harmless septics' may sound unfamiliar, but many villagers in the district are already using it. The three-chamber septics, as the name suggests, are composed of three interconnected sealed septic tanks. The faeces treated by this process are basically free of bacteria and parasites, so they can be called 'harmless'.

The key to building sanitary toilets and harmless septic tanks is to ensure the quality of construction, and the materials used must also meet national standards.

Detailed version

There are many types of rural sanitary toilets. In this district, we usually use flushing squat toilets. Human and livestock wastewater treatment generally goes through sewers, biogas digesters or three-chamber septics. However, some residents still use pit latrines or dig pits to bury their excrement.

The construction and renovation of sanitary toilets are not complicated and do not take much time. Professional workers and construction crews can renovate a sanitary toilet in half a day. Retrofitted sanitary toilets are more comfortable, more hygienic, and also safer. There is no need to worry about children or your mobile phone or other items falling into the dirty pit latrines. And the general flush toilet is very convenient, you can directly pour some water to clean up the faeces; you can also choose to install a flushing facility to clean up, just press it lightly, and the water will flow out and take away the wastewater. The management and maintenance of sanitary toilets are very simple. You only need to clean the toilets and deal with the stains in time when doing household work. If necessary, disinfectants can also be sprayed regularly.

In the meantime, we are actively promoting the construction of three-chamber septic systems. 'Three-chamber septic' may sound unfamiliar to you but many of your neighbours are using it already. As the name suggests, three-chamber septic is composed of three interconnected sealed septic tanks. The first tank is called the retention sedimentation and fermentation tank; the second tank is called the re-fermentation tank; the third tank is called the manure storage tank. The reason why it is called 'harmless' is because it utilises the principles of anaerobic fermentation. The bacteria and parasites in the faeces are killed in the first and second tanks. When the faeces flow into the third tank and are stored, it can already be called 'harmless'. The key to the construction of the three-chamber harmless septic tank is to ensure the quality of the building, and the installation quality of the manure pipe must be good. After using the manure stored in the third tank for fertilisation, it is crucial to pay attention to cover the septic tank tightly, and spray

some disinfectant if necessary to ensure safety and hygiene.

In short, the use of sanitary toilets and better sewage treatment facilities will produce no odour, no noise, no adverse effects on the surrounding environment, and it can achieve the goal of no secondary pollution.

3. **Benefits (BE)**

Brief version

According to 'Report on the Progress of the Toilet Revolution' released in 2017, 80% of infectious diseases in rural areas are caused by faecal pollution in toilets and insanitary drinking water. More than 30 types of infectious diseases are related to faeces, such as cholera, hepatitis, infectious diarrhoea, etc. The toilet revolution can not only make the countryside prettier, cleaner, and more convenient, but also effectively reduce the spread of diseases, enable rural residents to develop good hygiene habits, ensure their physical health, and accelerate rural civilisation.

Detailed version

Toilets are tightly related to people's livelihood. The toilet revolution aims to improve the rural environmental sanitation infrastructure and promote the rural civilisation. Maybe you will ask the question 'I have already started to use the sanitary toilet, so why should I participate in the toilet revolution?' Or maybe you will think about the relationship between the toilet revolution and 'me': 'I have used pit latrines for many years, and get used to it. I also don't think I need to upgrade my domestic wastewater treatment facility.'

These thoughts and confusion are quite normal. Next, we will give you a detailed explanation of the toilet revolution and the benefits it can bring you.

The first and most obvious benefit is the environmental benefit. Through extensive patriotic sanitation campaigns and continuous promotion and construction of sanitary toilets, the situation of no toilets or filthy toilets in the countryside has been changed. The contamination of water and soil by faeces has been effectively prevented. At the same time, the renovated toilet has no odour, and the density of mosquitoes and flies is reduced,

which improves the rural living environment greatly. The toilet revolution is beneficial to rural living conditions at both the household and the national level.

The three-chamber harmless septic system we are now promoting can further improve rural sanitation infrastructure. Human waste after its treatment is basically harmless. It will be more environmentally friendly to recycle the treated wastewater as fertiliser or discharge it directly into the sewer for unified treatment.

At the same time, upgrading toilets to cleaner ones can help eliminate faecal pollution, reduce the incidence of infectious diseases, and bring health benefits to us. In many rural areas of China, humans and animals live in the same area, and they are in frequent contact with each other. Many poultry faeces are treated in a correct way, resulting in the spread of pathogens in the faeces, and more zoonotic diseases, and this is a big concern to human health. Researchers have shown that effective toilet improvement can reduce children's diarrhoea by 36%, schistosomiasis infection rate by 77%, ascariasis infection rate by 29%, and child malnutrition by 36%.

As an inseparable part of China's rural revitalisation strategy, the rural toilet aims to promote rural civilisation and social progress. Promoting the 'toilet revolution' will help ensure rural residents' access to basic public health and sanitation services and narrow the gap in basic life infrastructures between urban and rural areas. Behind the rural toilet revolution is a healthier and more civilised life for villagers. As the living environment becomes more comfortable, the health condition becomes better and the air becomes cleaner, the happiness index of the people will gradually increase.

4. **Harms (HA)**

Brief version

Pit latrines have been used for many years, but they smell bad and breed mosquitoes and flies. Especially in summer, due to the increase of microorganisms, the mosquitoes and flies on the toilet are buzzing, the maggot eggs are squirming, the faeces ferment quickly, and bursts of pungent smells. The stench spreads with the wind, which adversely affects villagers' living environment and pollutes the air. In addition, mosquitoes and flies

carry bacterial viruses and spread among the population, leading to the occurrence and prevalence of various diseases such as bacillary dysentery and hepatitis A. Mosquitoes and flies that breed in dry toilets will contaminate food, causing food safety problems and threatening people's lives and health.

Detailed version

Dirty, messy, and smelly pit latrines that have been used for years in the countryside cause psychological discomfort, endanger residents' health, and reduce well-being and personal development. Improperly built pit latrines can also create serious safety hazards. According to reports, in May 2015, two elementary school students in Baotou, Inner Mongolia, drowned after falling into the dry toilet in the school. In February 2020, a 7-year-old boy fell into an open-air dry toilet in Nan'an, Fujian and drowned.

Because the excrement and urine in pit latrines cannot be removed in time, this kind of toilet has a foul smell. Especially in summer, the pungent smell will be especially obvious and bothering. There are also a large number of pathogenic microorganisms and parasite eggs in the faeces. If they are not treated in time, they will turn into flies, which increases the number of pathogens in the environment and spread pathogenic parasites. Mosquitoes and flies carry bacterial viruses and spread among the population, leading to the occurrence and prevalence of various infectious diseases such as bacillary dysentery and hepatitis A. Mosquitoes and flies that breed pit latrines will also contaminate food at will, causing food safety problems and threatening people's lives and health.

Pit latrines will also cause damage to the surrounding ecological environment and pollute groundwater resources. The leaching of faeces pollutes surface water and groundwater through surface runoff, making the water black and smelly and resulting in the death of fish and other organisms in the water. The accumulation and fermentation of faeces produce ammonia, hydrogen sulfide, methyl mercaptan, and other harmful gases, which not only endanger health but also seriously worsen air quality.

Furthermore, pit latrines and poor sanitation reduce well-being, hinder individual development—especially for children—and undermine sustainable regional development. The World

Health Organization has shown that poor hygiene is the direct cause of stunting in children, and when children's physical health is at risk, their attendance and academic performance will also be affected, further hindering children's access to education and their healthy development.

The rural toilet revolution is a sanitation campaign that aims to more effectively cover universal sanitation, improve the health of rural residents, and in turn promote economic growth, narrow the urban-rural gap, and further advance sustainable development.

5. **Ending**

If you have any questions about the toilet revolution or related toilet improvement projects, you can go to your local village committee office to get information and seek help. The responsibilities of the village committee mainly include handling the village's public affairs and public welfare undertakings, mediating disputes, assisting in maintaining social order, and promoting rural development.

Township and village cadres and comrades will actively answer your questions, help you clear up your confusion, and solve practical difficulties for you. At the village committee, you can also find information on other common poverty-alleviation programmes for farmers. If you have questions about poverty alleviation or agricultural support policies, or want to follow these policies more closely, the village committee is a reliable and timely source of information.

Finally, thank you so much for watching! Don't leave, everyone, and please take about 10 minutes to fill in a questionnaire. While you fill in the questionnaire, we will give you a small gift, thank you!

B.1 Video screenshots

Figure 5: Example 1

Figure 6: Example 2

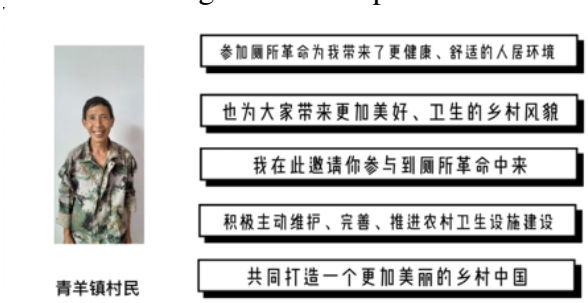

Figure 7: Example 3

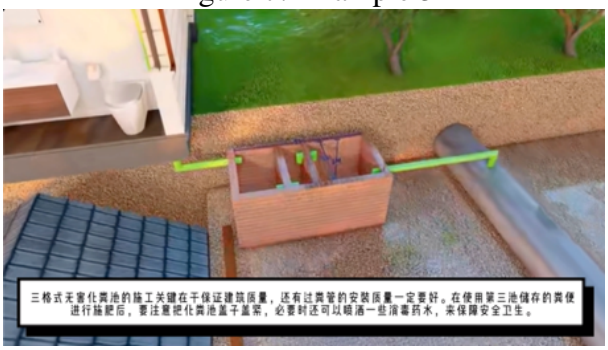

Figure 8: Example 4