How Mechanistic Explanations Reshape Learning and Behavior: Evidence from a Fertilizer Choice Experiment in Eastern Uganda

Anirudh Sankar[†] * Robert Dulin[‡] Benjamin Davies[†] Vesall Nourani[§] Jess Rudder[¶]
Abraham Salomon[‖] Godfrey Taulya[#]

May 3, 2026

CLICK HERE FOR LATEST VERSION

Abstract

Mechanistic explanations—descriptions of a system through the causal interactions of its parts—play a key role in human cognition and scientific progress. Despite their importance, we lack systematic evidence on whether and how mechanistic explanations help lay decision-makers interpret information in complex economic environments. We evaluate the causal impact of including mechanistic explanations in an information intervention: public demonstrations of fertilizer use for smallholder tomato farmers in Eastern Uganda. In all demonstrations, extension officers showcased the impact of a recommended fertilizer recipe. In the treatment group, officers also explained the mechanisms underlying the recipe's effects—introducing the language of macronutrients and the causal processes linking nutrients, soil features, and plant growth. We collected detailed data on beliefs and behaviors from 797 farmers in a lab-in-the-field experiment conducted at the demonstration site and followed up with them over two growing seasons. In the lab-in-the-field, treated farmers generalized more effectively—making better substitution and arbitrage decisions among fertilizers and achieving 9% higher simulated profits in an incentivized fertilizer-application task. At endline, treated farmers' real fertilizer choices reflected improved nutrient timing and balance, and their yields were 14% higher.

*Job Market Paper.

We are grateful to Agriworks Uganda for an incredible partnership—organizing the demonstrations, mobilizing farmers and providing for their comfort, connecting us with agronomists and extension officers, and more. We would particularly like to acknowledge the outstanding efforts of Boaz Otieno and Emma Wataka. We also thank the survey firms Matrice360 and VIE for maintaining high standards while nimbly navigating field challenges. We especially recognize our field and data managers, Henry Mulissa and Patrick Minja, for their close attention to data quality and strong on-the-ground judgment. We are very grateful to Sara Restrepo-Tamayo for her generous guidance on materials and logistics, and for coordinating the many moving parts of a challenging pilot. We thank ProsperPlay for developing a user-friendly digital fertilizer-application tool. Sankar thanks his advising committee for their invaluable guidance. The authors also thank colleagues at Princeton, Stanford, the University of Chicago, and beyond for helpful comments and suggestions. Financial support from the IGC (full research grant), the Stanford King Center (exploratory and advanced awards), SIEPR (George P. Shultz dissertation support), and the Weiss Fund (research grant) is gratefully acknowledged. Stanford IRB and UNCST approved the research protocols. This study was registered on October 28, 2024 (AEARCTR-0014649).

Affiliations: [†]Department of Economics, Stanford University; [‡]Department of Economics, Princeton University; [§]Harvard Graduate School of Education; [¶]Department of Applied Economics, Oregon State University; [‖]Agriworks Uganda; [#]School of Agricultural Sciences, Makerere University.

1 Introduction

To make informed choices, people need good data, but also good models to interpret the data. Indeed, a faulty mental model can lead to poor choices even when data are abundant (Gabaix, 2014; Schwartzstein, 2014; Spiegler, 2020). One class of model — central to both human cognition and scientific progress — is a mechanistic model: one that explains an observed result in terms of the causal mechanisms that produced it. In this paper, we ask whether providing mechanistic explanations of *why* an input produced an effect, over and above the data on the effect, improves learning and behavior in a real-world context.

Mechanisms are "entities (or parts) whose activities and interactions are organized so as to be responsible for [a] phenomenon" (Glennan, 2017). Mechanistic explanations go beyond statistical evidence of causal effects—for example, evidence that flipping a switch turns a light on or off—by describing the processes that produce the effect, such as the flow or interruption of current through a wire (Ahn and Kalish, 2000). The search for mechanisms is a central pursuit in science.[1] An influential literature in psychology and cognitive science suggests that mechanisms are also central to human cognition, arguing that from early childhood, people seek to learn mechanisms underlying the phenomena they encounter (Keil, 2022).[2]

We explore the causal impact of adding mechanistic explanations to an archetypal information intervention in a complex, high stakes domain. Specifically, we integrate mechanistic explanations into an agricultural demonstration, a frequently employed information intervention for smallholders that exhibits the impact of a novel technology or input by displaying a "trial" and control plot side-by-side, allowing the results to speak for themselves. We focus on fertilizers, a crucial but complex technology. Fertilizers are essentially food supplements for plants, and one applies some "recipe" of these synthetic colored pellets at various points in the crop growing season. As in cooking recipes, the efficacy of fertilizer recipes can be demonstrated solely through their results, as a black box. A mechanistic explanation would further describe why, biophysically, a fertilizer recipe had its result.

Prevailing approaches to government extension for smallholders de-emphasize mechanistic explanations (Anderson and Feder, 2007). A black-box approach to information interventions is also common in other complex domains, such as health or finance; these interventions demonstrate that a choice works through its outcomes (e.g., "this choice works well") or by referencing successful proxies (e.g., "Sara

[1] As (Machamer, Darden and Craver, 2000) argue in a now-classic paper: "In many fields of science what is taken to be a satisfactory explanation requires providing a description of a mechanism." A prominent strand in modern philosophy of science has argued that the practice of science itself is essentially the search for mechanisms (Craver, Tabery and Illari, 2024).

[2] Children ubiquitously ask "why" in search of mechanistic answers (Frazier, Gelman and Wellman, 2009), and even preverbal infants fixate on physically implausible events until a mechanism is revealed (Stahl and Feigenson, 2015; Perez and Feigenson, 2022). Likewise, most adults' first instinct in diagnosing events is to ask about mechanisms (e.g., for a car crash: "Were the brakes faulty?") rather than about statistics (e.g., "How often do accidents occur on this road?") (Ahn et al., 1995).

chose this, and she's like you"). Such emphases may reflect a perception that the complexity of these domains is simply too overwhelming for lay agents—after all, the mapping from choices to outcomes is indeed intricate—or a belief that mechanistic explanations, being more cognitively demanding, belong in formal schooling.[3] Yet simply unveiling a good choice may be inadequate when agents must adapt it to their own circumstances (Alidaee, 2023; Glover, Venot and Maat, 2017), or when they mistrust opaque recommendations (Jussupow, Benbasat and Heinzl, 2020) or find them unengaging. In many developing contexts, access to formal education may be limited, and any mechanisms once learned in school may be long forgotten. Mechanistic explanations may matter most precisely at the moment an agent faces a high-stakes decision in a complex environment—yet we lack systematic evidence on their real-world implications in such settings.

Mechanistic explanations may impact learning and behavior through two channels. The first is instrumental: mechanisms can aid generalization out of sample. In science, mechanistic explanations are uniquely powerful for generalizing from complex phenomena (Craver, Tabery and Illari, 2024). Indeed, a key aim of empirical economics is to identify mechanisms behind treatment effects, to assess what those effects imply for counterfactual policies or other populations (Banerjee and Duflo, 2009; Deaton, 2010). Mechanistic explanations have also helped laypersons generalize. For example, the basic idea of germ theory—that invisible germs cause illness—continually guides common-sense precautions in new circumstances, from handwashing to avoiding stagnant water.[4] A mechanistic understanding sharpens a learner's priors and posteriors out of sample, which can lead to better choices when the learner's problem also extends beyond the sample context. The second channel through which mechanistic explanations may matter is behavioral: for example, sharing mechanisms may increase engagement or make the explainer more credible.[5] Experimentally adding mechanistic explanations to an information intervention engages both channels, but we will especially test for the instrumental channel—to assess whether learners actually apply this conceptual model to decisions across contexts.

We conceptualize mechanistic explanations of a phenomenon as comprising a *language* and a *structure*—both deriving from scientific knowledge—and show that each enables different forms of generalization. The *language* names conceptual parts of a system and relates otherwise disparate choices. Fertilizers, for example, are mechanistically described through the *language of macronutrients*: to a first order, fer-

[3]In qualitative interviews with both extension workers and farmers in the study region, open-ended questions about what farmers wanted to know in a demonstration consistently elicited the ambiguous response that they wanted to know "the results." Interpreted one way, this orientation towards results may have led extension workers to view mechanisms as peripheral to farmers' concerns.

[4]Conversely, before the emergence of germ theory in the mid- to late-19th century, routine preventive hygiene—such as handwashing in medical settings—was not widespread (Matthias, 2025).

[5]These "behavioral" benefits may emerge from improved generalization or from other causes, including innate cognitive preferences. For example, children both enjoy mechanistic explanations and readily generalize from them (Keil, 2022), but it is unclear whether they enjoy it *because* they expect to be able to generalize from them. For simplicity, we will call any benefit that is not the literal act of generalization "behavioral."

tilizer action is mediated by the concentrations of three macronutrients. Providing a conceptual language amounts to providing a form of dimension reduction that makes any set of observed data or "correlations" more informative (Davies and Sankar, 2025). After naming conceptual parts, mechanistic explanations then describe how the interactions among those parts are causally structured to produce some effect. Fertilizer action is mechanistically explained by the *structure of nutrient response*: knowledge of how these macronutrients move through soil given observable characteristics, and how they contribute to plant growth. Even qualitatively described nutrient-response structures can be used to anticipate effective timing of macronutrient applications and guide diagnostic experimentation for identifying an optimal configuration (Davies and Sankar, 2025). They also guide how recipes should be adapted across soil contexts—an important ability in environments with pervasive heterogeneity in agricultural conditions across space and time (Suri and Udry, 2022).

Our specific intervention is embedded within fertilizer demonstrations for tomato farmers in Eastern Uganda, organized by an NGO and led by government extension officers.[6] These smallholder farmers are far from the production frontier, fertilizer is an important but inefficiently adopted technology, and in this context, extension-led demonstrations are a typical mode of technology promotion.

We randomize the inclusion of mechanistic explanations of fertilizer in demonstrations, just before the upcoming growing season, to study their causal impact. In all demonstrations an extension officer showcased the impact of a fixed fertilizer recipe on a real plot, side-by-side with a control plot. During treatment demonstrations, the officer also explained why it worked—covering its nutrients, their interactions in soil, and role in plant growth. During control demonstrations, the officer simply walked through the recipe and discussed the results, without providing explanations of the mechanisms behind its impact. He also delivered a placebo training on crop management unrelated to fertilizer science.

We leverage survey data from two phases: a lab-in-the-field phase at the demonstration site itself, and an endline phase one and two seasons following the demonstration. The lab-in-the-field data includes fine-grained measures of farmers' beliefs about the fertilizer production function using a range of choice problems—substitution, arbitrage, counterfactual recipe predictions, translation across soil contexts, willingness to pay, responses to marketing claims, beliefs about broader meta-principles, and, finally, adaptive experimentation in a sandbox module. Many of these measures are incentivized and directly comparable to agronomic benchmarks.

The endline captures changes to real-world practices and outcomes, and evolution in knowledge and curiosity. We re-interviewed 92% of farmers (75% of whom grew tomato plots in one of the two growing seasons following the demonstration). Anticipating possible decay in recall of mechanistic details (Fisher

[6]Indeed, our partner organization, Agriworks, frequently conducts these demonstrations on various agricultural practices, especially irrigation.

and Keil, 2016), after the first growing season we invited a random half of treated farmers to a brief reinforcement training that reprised the fertilizer science. This session was held indoors (i.e., off-farm), introduced no new material, and allowed us to test whether repetition of mechanistic explanations was necessary to sustain mechanistic understanding.

For our lab-in-the-field phase, many of our tasks are designed to mimic real-world problems encountered by farmers in the course of their daily activities, such as how to choose a substitute for a fertilizer that is not available at the local agrodealer, or to predict the yield they would achieve from applying a novel fertilizer recipe. Importantly, these tasks do not reward regurgitating facts from the training, but necessitate generalization. Improved farmer performance on these tasks, then, is suggestive not only of increases in knowledge, but of an enhanced ability to deploy this knowledge in the relevant economic environment. We use the conceptual framework to categorize treatment effects on these choice problems into *language* and *structure* tests, depending on which component of mechanistic explanations farmers must draw on to arrive at agronomically appropriate decisions.

Our *language* tests show that treated farmers are more likely to recognize macronutrients as the relevant "active ingredients" behind fertilizer. Treated farmers are better able to use knowledge of macronutrients to substitute fertilizers and exploit arbitrage opportunities. The need to substitute fertilizers or make "arbitrage" decisions between available fertilizers is a common experience in this region.[7] However, treatment does not induce farmers to fully collapse choices into macronutrient equivalent classes; farmers consider macronutrients only as partially sufficient: they order choices but may not tell the whole story of fertilizer action.

The *structure* tests reveal that treated farmers adjust their decisions to account for both timing and diminishing returns to macronutrient supplementation, tailoring their fertilizer use to context-specific differences in effectiveness. This is most evident in the adaptive experimentation module, where treated farmers achieve 9% higher simulated profits by configuring more agronomically viable fertilizer recipes, across two tested soil types (the soil used in the demonstration and a counterfactual soil with different observable properties), without increasing costs. This attentiveness to high and diminishing marginal returns is also confirmed in separate tests of operative meta-principles and beliefs about how best to augment a fertilizer recipe.

The endline results indicate that treatment affected fertilizer input choices in ways consistent with sound agronomic reasoning. Treated farmers' applications of nitrogen and potassium were better balanced and well-timed respectively, the latter (potassium at fruiting) being a key mediator of profits in the adaptive experiment. Turning to downstream outcomes, treated farmers achieved 14% higher yields

[7] 26% of farmers in our sample report not being able to find a fertilizer they wished to purchase at their local agrodealer in the past 12 months, and 52% report finding that the price of a fertilizer they wished to purchase was unexpectedly high.

per acre. There is also a positive but imprecisely estimated effect on profits. The marginal effect of the reinforcement training on the treatment-affected outcomes—over and above the main treatment—is indistinguishable from zero, suggesting that any decay in mechanistic understanding likely occurred quickly after an explanation rather than gradually over a growing season.

This paper contributes to three strands of literature. First, we contribute to a growing empirical literature on mental models (Esponda, Vespa and Yuksel, 2024; Hanna, Mullainathan and Schwartzstein, 2014; Laajaj and Macours, 2024; Miehe et al., 2025). Much of this work also studies smallholder agriculture—an environment where lower-income farmers, often with limited formal education, make high-stakes decisions under quintessential complexity (De Janvry, Macours and Sadoulet, 2017). These papers typically focus on specific defects in agents' mental models, revealed through systematically poor choices, and intervene to overturn those choices through black-box treatments. In contrast, our focus is on the intervention itself: we share a causal model with farmers and study its impact on learning and behavior.

Second, we contribute to research on the value of conceptual learning. A long literature in education, philosophy, and psychology contrasts learning conceptually with learning by rote (Dewey, 1927; Haeusler and Donovan, 2020; Keil, 2022; Nourani, Ashraf and Banerjee, 2025; Vygotsky, 1978). Our conceptual framework draws on theoretical work (Davies and Sankar, 2025) that quantifies the value of conceptual knowledge—deeper understanding of *true* models—within a statistical decision-theoretic framework.[8] We experimentally validate the theory's high-level prediction that, in farmers' "small-data" world, conceptual knowledge improves the interpretation of exogenous information and guides the collection of diagnostic information. This theoretical framework deliberately rules out the pervasive model misspecification of "models as narratives" approaches (Eliaz, Galperti and Spiegler, 2025; Schwartzstein and Sunderam, 2021), motivated by the observation that in many domains science has converged on objectively true models that agents recognize as ones that can be learned. This perspective allows us to treat conceptual understanding as an aspect of human capital (Mokyr, 2016; Sen, 1999).

Finally, we contribute to the literature on information frictions in technology adoption, which emphasizes that one dimension of under-adoption of productivity-enhancing technologies is informational (Magruder, 2018). Farmers frequently discount information signals in heterogeneous environments, suspecting that what works for another farmer—or on a demonstration plot—may not work for them (Alidaee, 2023; Chakraborty, 2022; Munshi, 2004; Tjernström et al., 2021). This concern is often justified: in traditional ecologies, returns to many agricultural technologies vary sharply over small spatial scales

[8]For example, a class of "true" models such as (Callander, 2011; Callander and Matouschek, 2019) employs a conservative heuristic: payoffs vary smoothly for nearby choices but are uncorrelated for distant ones. These models effectively refuse to sparsify, implicitly maintaining high-dimensional representations to avoid misspecification. Conceptual knowledge allows learners to move beyond such locally smooth but "long" representations by leveraging latent structure that governs uncertainty in payoffs.

as soil properties, for example, shift rapidly (Suri and Udry, 2022). We ask whether sharing mechanistic explanations helps farmers adapt and translate information signals rather than discount them altogether.

2 Conceptual Framework

We isolate *language* and *structure* as the two essential components of mechanistic explanations, that interpret complex phenomena and enable distinct forms of generalization. The language names important parts of a system and highlights which choices are expected to yield similar outcomes. The structure describes how the named parts work together, helping predict specific outcomes. For fertilizers, the *language of macronutrients* and the *structure of nutrient response* together explain the high yields from sophisticated fertilizer applications and guide the adaptation of fertilizer recipes across contexts.

Even for a single crop like tomatoes, holding other inputs fixed, fertilizer application is a complex decision. Fertilizers can conceivably be applied at any point in an 8–10 week growing cycle. There are dozens of fertilizers in the market, a continuum of application quantities, and returns to any application depend on (1) timing within the schedule, (2) other fertilizers applied at the same or different times, and (3) soil properties. The action space is combinatorially massive. Despite this, some promising fertilizer "recipes" have been determined. A potent and cost-effective recipe for the "infertile" soil shown in Figure 1a is shown in Figure 1b. This is exactly the agronomist-formulated recipe we demonstrate experimentally, on a soil archetype of Figure 1a. At multiple points in the schedule, two or more fertilizers are mixed, with quantities varying across application points. Without a model to reason about fertilizer application, such a recipe is a needle in a haystack.

A mechanistic explanation rationalizes this recipe at two levels. The *language of macronutrients* says that fertilizer action is first-order mediated by its concentration of three macronutrients—nitrogen (N), phosphorus (P), and potassium (K). Table 2a (from Figure 2) shows macronutrient composition for some common fertilizers, which are typically prominently printed on bags as in Figure 2b. What really matters in a fertilizer recipe, then, is the corresponding distribution of macronutrient supplementation. Figure 1c shows the macronutrient distribution for the recipe of Figure 1b. In practice, this distribution can be chosen first, with the concrete fertilizer instantiations resolved in a "second stage" as a purely mathematical problem. Reasoning through a macronutrient distribution requires abstraction, since macronutrients are not directly observable in either fertilizers or soil. The language of macronutrients is therefore *conceptual*. Naming conceptual parts before describing their interactions is constitutive of many scientific mechanisms (Craver, Tabery and Illari, 2024).[9]

[9]Reasoning in a conceptual space of macronutrients goes beyond the general idea that fertilizers provide nutrients. Some studies on smallholder fertilizer decisions portray farmers as reasoning with specific nutrients, though the role of nutrients in their thinking is unclear. For example, (Islam and Beg, 2021) use leaf color charts to address nitrogen deficiencies or over-

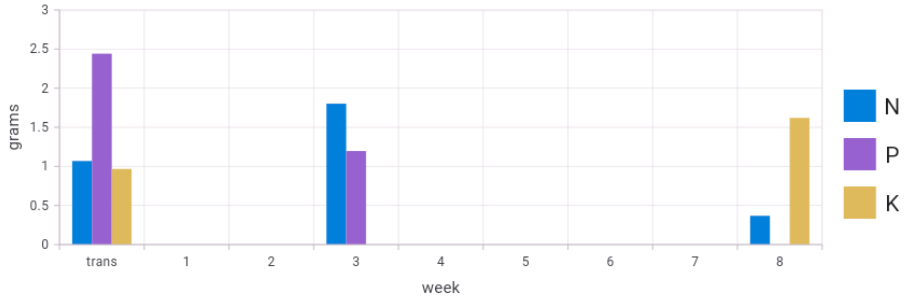


(a) Grey Soil

Week	Fertilizer 1	Fertilizer 2	Fertilizer 3
Transplanting	NPK 17:17:17, 1 glass cap	DAP, 3 glass caps	MOP, 1 glass cap
Week 3	DAP, 1.5 glass caps	Urea, 2 glass caps	
Week 8	MOP, 1.5 glass caps	Urea, 1 glass cap	

Quantities are per-plant in amounts (bottle caps) familiar to farmers.

(b) Agronomist-formulated Fertilizer Recipe for Grey Soil



(c) Macronutrient distribution of Fertilizer Recipe

Figure 1: Recipe and Nutrient Distribution for Grey Soil

Fertilizer	Percent N	Percent P (P_2O_5)	Percent K (K_2O)
Urea	46	0	0
DAP	18	46	0
MOP	0	0	60
CAN	27	0	0
TSP	0	46	0
NPK 25:5:5	25	5	5
NPK 17:17:17	17	17	17

This includes generic fertilizers commonly available in Uganda. There are many more branded formulations, all of which have to label their nutrient compositions (Ministry of Agriculture, Animal Industry and Fisheries, Republic of Uganda, 2016).

(a) Macronutrient Composition of Common Fertilizers



(b) Bag of DAP fertilizer. The macronutrient N-P-K composition 18-46-0 is printed clearly

Figure 2: Fertilizer nutrient information presented in table (a) and packaging example (b).

The instrumental value of a conceptual language is that it relates seemingly disparate choices. For the problem of fertilizer application, there is a complicated payoff (yields or final profits) function $f(x)$ of choices (fertilizer recipes) $x \in X$. The language refers to a space Z (macronutrient distributions) where $dim(Z) < dim(X)$, and a dimension-reducing function $z(x)$ such that there exists a $g(z)$ with $f(x) \approx g(z(x))$. The language thus rewrites the problem up to a global approximation, as in Figure 3. In (Davies and Sankar, 2025), we show that that this language frequently makes any set of data or "correlations" more informative. The basic intuition is that a single outcome of $f(x')$ is informative about the entire equivalence class $\{x|z(x) = z(x')\}$. Per Table 2a, macronutrient equivalence classes are often infinitely sized, so the outcome of single recipe can be informative of infinitely many recipes.[10]

supplementation, but these charts merely map leaf colors to Urea applications, without requiring reasoning about nitrogen itself. Likewise, (Bold et al., 2017) and (Hoel et al., 2024) claim to elicit beliefs about the nitrogen content of fertilizer bags to study perceptions of quality degradation, but in practice 'nitrogen' in these elicitations functions more as a proxy for overall fertilizer potency (its "oomph").

[10]This dimension reduction is lossy: for example, different fertilizers release nutrients at different rates depending on the chemical compound. (Davies and Sankar, 2025) assumes agents know such conceptual representations are incomplete and shows that agents optimally "regularize" to a depth threshold rather than continually "deepening" to finer distinctions. Consistent with this, agronomic nutrient-response studies typically abstract from fertilizer formulation despite known release differences.

The *structure of nutrient response* describes how the three nutrients move in soil, depending on observable soil properties; how they participate in different stages of plant growth; and how they interact. By deducing implications from these descriptions, learners refine their priors and posteriors over the marginal returns to supplementation as functions of timing, observable soil properties, and nutrient interactions. In the formulation $f(x) \approx g(z) \circ z(x)$ (Figure 3), the *structure* informs $g(z)$ itself.

The *structure* rationalizes the general shape of a potent nutrient distribution like Figure 1c. For example, nitrogen is important throughout tomato growth but especially during vegetative growth. In this soil with a lower percentage of clays in the mineral fraction (Figure 1a), nitrogen is more easily leached away during irrigation or precipitation, and must be resupplied in order to maintain the same nitrogen levels in the soil solution. This is why, in Figure 1c, nitrogen is applied repeatedly ("split application") with particular emphasis in week 3. Since each application of fertilizer is labor-intensive, one can also experimentally vary the number of nitrogen splits to examine where returns diminish. More generally, understanding this structure helps learners design diagnostic experiments to identify an optimal fertilizer configuration (Davies and Sankar, 2025).

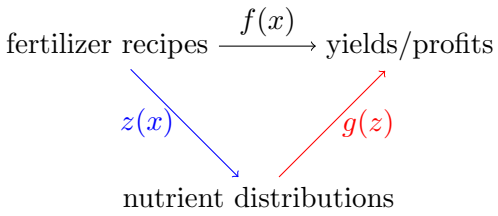


Figure 3: The *language of macronutrients* (blue) and the *structure of nutrient response* (red) jointly explain a complex payoff function $f(x)$ over fertilizer recipes. The language provides a known mapping $z(x)$ that rewrites $f(x) \approx g(z) \circ z(x)$, while the structure conveys information about $g(z)$ itself.

A particularly important generalization using the structure of nutrient response is across soil contexts, since in these traditional ecologies soils rapidly change in ways that substantially impact fertilizer returns (Suri and Udry, 2022). Here, the structures provide important clues for adaptation. For example, the black soil archetype Figure 4a contains higher clay content, and has more chemically active types of clay than the soils in 1a. These clays provide higher capacity for the soil to retain nitrogen, reducing leaching. As a result, to achieve the same yield as the recipe for the "infertile" soil Figure 1b, both the amount of nitrogen and the number of split applications can be reduced, as in 4c (with example recipe instantiation 4b). Helping farmers translate what they observe across context could help mitigate a major impediment to farmers' learning under heterogeneous conditions (Alidaee, 2023; Chakraborty, 2022; Munshi, 2004; Tjernström et al., 2021).

The full qualitative mechanistic explanations of fertilizer (our treatment) are provided in Appendix B. Mechanistic explanations can vary in depth, and the appendix clarifies the level of mechanistic detail used



Perceivable characteristics:

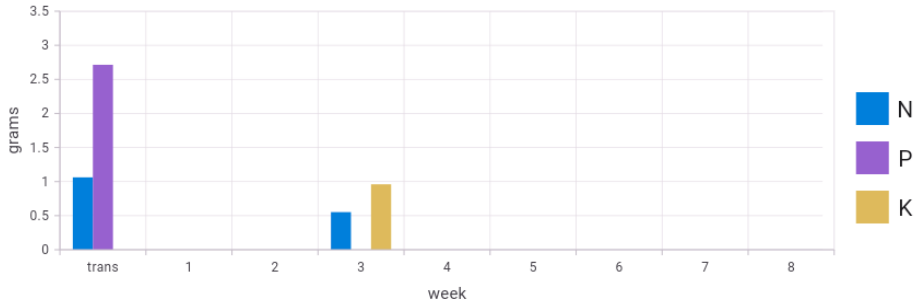
- black
- cracking clays
- very sticky when wet
- Very fertile
- Takes water slowly but it holds for long (crop remains healthy even after a medium drought)

(a) Black Soil

Week	Fertilizer 1	Fertilizer 2
Transplanting	DAP, 2 glass caps	
Week 3	Urea, 0.5 glass caps	MOP, 0.5 glass caps

Quantities are per-plant in amounts (bottle caps) familiar to farmers.

(b) Agronomist-formulated Fertilizer Recipe for Black Soil



(c) Macronutrient distribution of Fertilizer Recipe

Figure 4: Recipe and Nutrient Distribution for Black Soil

in our intervention. For instance, we do not explain why potassium matters for fruit filling in terms of enzyme activation; our descriptions of nutrient roles in plant growth (e.g., "potassium supports the filling of fruits, grains, and tubers") are closer to teleological accounts (Joo, Yousif and Keil, 2021). By contrast, the explanations of how nutrients move in soil have greater mechanistic depth, though still reasonably bounded. For example, we state that clays hold nitrogen better but stop short of discussing the electro-

statics underlying clay–nitrogen retention. As implemented, these qualitative mechanistic explanations already permit key forms of generalization.

3 Experiment

We randomized the inclusion of mechanistic explanations as part of fertilizer demonstrations for Ugandan tomato farmers, with data collected from both lab-in-the-field activities and household surveys. This allows us to identify the causal impact of mechanistic explanations on both input beliefs and input choice, in a setting where the returns to appropriate inputs are high.

3.1 Setting

Our setting is well-suited for studying the role of mechanistic explanations in several respects. First, smallholder agriculture in Sub-Saharan Africa is a setting with persistently low productivity and substantial information gaps. Second, the technology we study—fertilizers—is a key but underadopted productivity-improving input. It is complex to deploy optimally; adoption is not a simple binary decision. Third, we embed our experiment in a canonical mode of technology dissemination—the agricultural demonstration—that is among the most common in this context and familiar to our participants.

Small-scale agriculture is the dominant sector in Uganda, comprising 66% of total employment (World Bank, 2025), and is prominent in our study region (Edaku, 2017). Tomato is an important vegetable crop nationally (Atuhaire, Ocan and Jørs, 2016), yet yields are estimated at only 39% of the frontier potential (Wacal et al., 2023). This mirrors the broader pattern in Sub-Saharan Africa, where small-scale agriculture is similarly dominant but remains far from the production frontier (Suri and Udry, 2022). Fertilizer use is very low nationally (Amann et al., 2021), as in many other African countries. While usage is somewhat higher for vegetable crops, it is still far from optimal: in our sample, even though the vast majority of farmers report having used inorganic fertilizer, less than 10% of baseline applications meet the basic agronomic guideline of supplementing all three macronutrients at some point during the growing season.

Our premise is that this suboptimal deployment is not only due to financial constraints (credit or liquidity) or limited market availability, but also by a binding informational friction: specifically, how to configure fertilizers appropriately and effectively. Indeed, relaxing financial constraints, including through fertilizer subsidies, has often had muted effects (Poulton, Kydd and Dorward, 2006; Karlan et al., 2014). Similar informational frictions have been documented in nearby Western Kenya (Chandrasekhar et al., 2022). A key challenge is the need for tailored information: in traditional farming ecologies, soil characteristics can vary at scales as small as tens of metres (Hengl et al., 2021; Suri and Udry, 2022), limiting the usefulness of knowledge externalities. The widespread paradigm of precision agriculture is, in

large part, a response to this formidable and highly contextual information problem (Nyaga et al., 2021).

Agricultural demonstrations and field days are core services provided by government extension systems in Uganda and across Sub-Saharan Africa (Ministry of Agriculture, Animal Industry and Fisheries Uganda, 2016; Agwu et al., 2023). These demonstrations are typically held on a designated farmer's plot (Buyinza et al., 2015), but with national extension-to-farmer ratios of roughly 1:2400 (Sserunjogi et al., 2023), they are rarely personalized: many farmers from the surrounding area are invited to attend. Demonstrations usually feature a successful "trial" plot alongside a control plot to let the results speak for themselves, a form of A/B testing with a long history in traditional farming (Hansson, 2019). However, the impacts of such demonstrations can be muted (Maertens, Michelson and Nourani, 2021); one reason may be that the demonstration context differs from farmers' own contexts (Alidaee, 2023), requiring adaptation that farmers may not know how to execute.

Our sampling frame of tomato farmers invited to the demonstrations was drawn from past or potential users of an on-demand irrigation service, Agriworks Uganda. Farmers were dispersed across the study region. Potential users were required to have prior irrigation experience, while past users of Agriworks had this by definition. This inclusion criterion of irrigation experience is not innocuous, since many farmers in the region and nationally rely solely on rain-fed production (Kizza et al., 2017; CultivAid, 2021). However, it offers two main advantages for our purposes. First, it allows us to hold irrigation—a complementary input—constant and make it common knowledge, minimizing ambiguity so that the intervention targets fertilizer knowledge specifically. Second, past irrigation experience makes these farmers more likely to produce in the dry season (when prices are higher) and suggests more careful consideration of crop inputs, increasing the stakes of the demonstration. The sampling frame included 1,269 farmers, with 49% prior users of Agriworks and 51% potential users. The latter were deliberately concentrated in the southern part of the study region, where Agriworks had not yet expanded but was considering future entry.

3.2 Intervention

We randomized the inclusion of mechanistic explanations in fertilizer demonstrations held at the outset of the growing season, to study their causal impact. In all demonstrations, an extension officer showed the impact of a fertilizer recipe on real plots. On treatment days, the officer also explained why it worked—covering its nutrient composition, the interactions of those nutrients in soil, and their role in plant growth. On control days, the officer instead delivered a placebo training unrelated to fertilizer science.

Every demonstration revealed the "reduced-form" causal impact of a fertilizer recipe on a particular soil

type by displaying eighth-acre "trial" and control tomato plots side-by-side, where the only difference in input application was the inclusion versus exclusion of the fertilizer recipe. The plots were shown at harvest stage, allowing the results to speak for themselves. The fertilizer application followed exactly the recipe in Figure 1b, and the soils corresponded to the grey soil archetype in Figure 1a. The recipe, formulated by our soil scientist co-author, was designed to be impressive in producing a moderate, attainable yield on this "infertile" sandy soil using relatively meagre amounts of fertilizer. The extension officer walked participants through the recipe in detail, bringing physical bags of all the fertilizers in Table 2a (from Figure 2) for close inspection. He also described the complementary inputs (constant between the trial and control plot) —hybrid seed (the most popular variety in the region), irrigation, pesticide use, staking, and weeding practices.

On treatment days, the officer further rationalized why the recipe achieved its observed effect, as described in Section 2. The mechanistic training builds a qualitative conceptual model of soil and nutrients so farmers can reason about fertilizer use using observable soil characteristics together with underlying soil and nutrient processes. It begins with soil composition and texture—especially how sand versus clay governs the soil's ability to retain water and nutrients—and then introduces the three key macronutrients (N, P, K), their roles in plant growth, and how each behaves differently in the soil. It links visible soil features to nutrient retention and availability, and emphasizes that plant growth is constrained by the scarcest required nutrient. It then elaborates how these mechanisms inform a decision framework for managing fertilizers along three margins—type, amount, and timing—and shows how this framework guides the design of the demonstrated fertilizer recipe for the specific soil conditions to which it is applied. The full treatment curriculum on fertilizer and nutrient science—developed with our soil scientist co-author to balance scientific rigor with farmer comprehension[11]—is provided in Appendix B.

Our collaborating extension officers received dedicated training on this curriculum and had prior experience presenting it in both a pilot and a soft launch. The fertilizer theory component of the treatment took approximately 1.5 hours. On control days, extension officers instead delivered a placebo training on other aspects of crop management (e.g., weeding, staking, pesticide use) to keep the total training time roughly equal (about 2 hours) across treatment and control groups.[12]

We held a location-stratified RCT at the farmer level, where the strata are eight broad geographical zones in the study region. We ran 18 demonstration days, arranged as 9 pairs ("tranches") of consecutive

[11]For example, Liebig's law of the minimum, which states that a plant's growth is determined by its most limiting (bottleneck) nutrient, is conveyed through the analogy of a Ugandan who, as much as he likes the maize staple *posho*, will not eat more once it is no longer complemented by some sauce.

[12]We included a placebo training to address concerns that exposure length itself may drive effects (Zajonc, 1968) We chose an agriculture-related training (an unrelated one would have been bizarre to farmers), even though this risked attenuating treatment effects through partial exposure. Because the lab-in-the-field exercises were focused narrowly on fertilizer beliefs, they were unlikely to have been affected.

days, with one randomly selected day in each pair assigned to treatment and the other to control. Each day had two identical demonstration sessions (morning and afternoon). Each farmer from the sampling frame was invited to exactly one session to receive their assigned treatment. Assignment proceeded in three steps: farmers were first randomized to treatment or control within location strata, then assigned to a tranche, and finally to a session between morning and afternoon. The comparison of invited farmers in the same session slot across the two days of a tranche forms an internally valid, location-balanced sub-RCT, even though the mix of districts varied across sessions and tranches by design to accommodate logistical needs.[13] All estimates are pooled across the full RCT with appropriate fixed effects. About 30–35 farmers were invited to each session, with 20–25 attending. Because non-attendees are unobserved, the analysis uses attendees only and checks baseline balance to rule out systematic selection into the treatment or control day.

Insofar as mechanistic explanations have any immediate impact on beliefs, their longer-term effects are uncertain. Psychological evidence suggests that recall of mechanistic *details* can decay rapidly over time (Fisher and Keil, 2016; Rozenblit and Keil, 2002). However, these studies come from settings where it is unclear how useful such details are to daily life.[14] In our context, where practical application is immediate, the durability of mechanistic detail is an open question. In principle, mechanistic understanding may even support retention by acting as a form of cognitive compression—allowing fertilizer practices to be rederived "from first principles." Nevertheless, to test whether any decay occurs, we implemented a reinforcement intervention between the first and second growing seasons. A random half of treated farmers were invited to a brief refresher session that revisited the fertilizer science. The training was conducted indoors (off-farm), introduced no new material, and allowed us to test whether repeating mechanistic explanations was necessary to sustain mechanistic understanding.

3.3 Lab-in-the-field data

We conducted a lab-in-the-field experiment with the 797 farmers who attended the demonstration site. The core modules elicit beliefs, opinions, and practices related to fertilizer use, and many measures are incentivized. Our approach avoids direct comprehension tests (e.g., asking "why does nitrogen matter?") and instead poses practical decision problems from which mechanistic understanding can be inferred. This is for three reasons. First, our emphasis is on knowledge that is instrumentally relevant rather than knowledge for its own sake. Second, direct knowledge questions risk alienating control respondents by

[13]For example, mudslides before the demonstration days made it easier to mobilize farmers from nearby locations in the first tranches. We also scheduled distant locations in afternoon slots to account for commute times.

[14]For example, Fisher and Keil (2016) show that college-educated adults often fail to recall details from their former major—such as biology graduates forgetting the Krebs cycle. Unlike farmers applying fertilizers, it is unclear how much these former majors needed the material outside of exams.

asking about material they were not exposed to. Third, this approach minimizes experimenter demand effects by avoiding explicit references to the training curriculum.

3.3.1 Structure of the lab-in-the-field

The lab-in-the-field modules were administered at the demonstration site immediately before and after each demonstration. Our pre-demonstration module allows us to measure baseline characteristics and pre-treatment beliefs and knowledge, while the surveys after the demonstration allow us to examine treatment effects. While the demonstration itself was administered to a group of farmers, and often involved group participation, all surveys were administered privately and simultaneously between each farmer and their designated enumerator, enabling us to capture beliefs in the absence of social image concerns or learning spillovers.

The sequence of lab-in-the-field activities is displayed in Figure 5. It begins with a baseline module on demographics, farm characteristics, and agricultural practices, followed by elicitation of priors about the yields of certain pre-specified fertilizer recipes. Farmers then attended their demonstration (along with treatment or control training), after which we elicited their posterior beliefs on the same pre-specified fertilizer recipes. This was followed by other belief and opinion modules, then an elicitation of willingness to pay for a novel potassium fertilizer included in the recipe, and finally an interactive adaptive experimentation game on a digital tool.

The activities in red are incentivized. For three of these activities—prior and posterior fixed-recipe prediction problems, and the interactive adaptive experimentation game—farmer responses were directly compared to our agronomist co-author's predictions, which formed the basis for incentives. The agronomist generated yield predictions for a factorial combination of nutrient quantities and timings, which we used to train a simple model $\hat{f}(x)$ for returns to fertilizer application (see Appendix C). For these three exercises, farmers were further randomized to one of two soil types: the demonstration soil (grey soil) or a counterfactual soil (black soil). Farmers were told which soil applied to them (with the counterfactual soil physically present for inspection) and made fertilizer assessments accordingly. Their fertilizer assessments were made in known complementary input conditions (in fact the very same complementary input conditions of the demonstration) that were common across all farmers [15]. This design allows us to test the translation of signals across soil contexts (Suri and Udry, 2022).

After the lab-in-the-field exercises, each farmer received a pamphlet that summarized the content of the training they received (control or treatment), so they could refer back to it in the future.

[15] The complementary input conditions were specified to be identical between grey and black soils, except that on the black soil it can be assumed that drainage canals were also used to prevent waterlogging. Respondents who were assigned the counterfactual black soil were informed accordingly

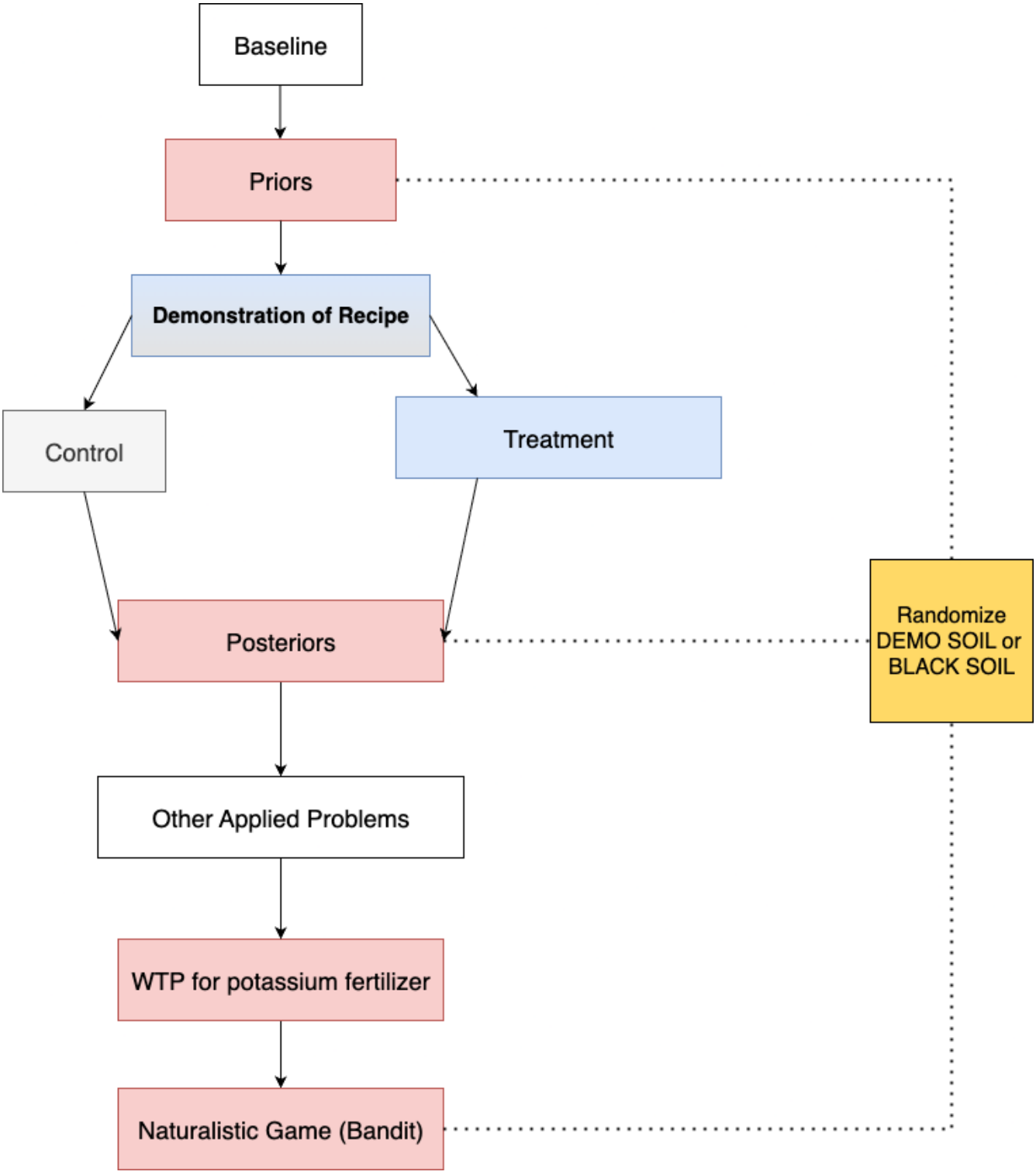


Figure 5: Lab-In-The-Field Activities

3.3.2 Baseline

A baseline survey was administered as soon as the invited farmer arrived, prior to the start of the demonstration. It established a comprehensive pre-intervention profile of farmer characteristics, agricultural production practices (especially for tomato growing), and fertilizer knowledge. Because non-attendees are unobserved, we rely on a joint baseline balance test to rule out differential selection into treatment versus control demonstration days within a tranche.

Farmer characteristics included a standard slate of socio-economic variables—age, gender, household composition, education, occupations, and farming history—along with a bar graph literacy test (relevant given frequent references to bar charts during training and lab-in-the-field exercises) and an unincentivized risk-aversion elicitation. We also collected detailed information on the structure of each farmer's production system, including farm layout, land allocation, and crop mix. Tomato production data covered varieties grown, seasonal patterns, and management practices, especially plot-level fertilizer use: types, quantities, timing, and methods. The instrument concluded with structured questions on fertilizer knowledge, particularly around soil fertility, nutrient roles, and fertilizer practices.

Summary statistics for a subset of baseline variables for the $N = 797$ demonstration attendees are shown in Table 6. The sample is predominantly male (83%), with an average of 18 years of farming experience, and earns about 71% of household income from farming. Farmers own an average of 3.2 acres, with 1.4 tomato plots. Although they grow a variety of crops, tomatoes contribute 56% of farming income. Plots are distributed across three main soil archetypes: black (69.8%), brown (17.8%), and grey (22.0%). These archetypes are not formal soil classifications, but rather heuristic categories that capture highly observable distinctions in texture and mineralogy where optimal fertilizer use is likely to differ. Only 27.5% of the sample have completed high school, though 59% have had some form of agricultural vocational training—suggesting the potential for richer explanations to be impactful precisely at the moment of information delivery. Baseline fertilizer knowledge is low: on average, farmers correctly identify the nutrients contained in just 6.7% of common fertilizers. For their self-reported fertilizer practices on their tomato plots, fewer than 10% meet the basic agronomic guideline of supplementing each macronutrient at some point in the growing schedule.

A joint test of balance between treated and control farmers yields a p-value of 0.459, so we fail to reject balance. Assignment to treatment did not increase attendance ($p = 0.687$), ruling out differential attrition.[16]

[16]Indeed, all demonstration invitees only knew that they were invited to fertilizer and agricultural demonstrations, not whether they were in the treatment or control informational variants.

	mean	sd	min	max	count
Male	0.831	0.375	0.000	1.000	752
High School	0.275	0.447	0.000	1.000	752
Years Farming	18.503	10.991	1.000	60.000	739
Pct Income Farming	70.949	20.436	5.000	100.000	739
Farm Size	3.229	4.814	0.125	90.000	739
Years Farm Tomatoes	9.995	8.275	0.000	51.000	739
Pct Farm Income from Tomatoes	56.243	24.170	0.000	100.000	738
# Tomato Plots	1.421	0.804	1.000	8.000	738
Grey Soil	0.220	0.414	0.000	1.000	738
Black Soil	0.698	0.460	0.000	1.000	738
Yield	5.468	5.515	0.180	25.200	738
# Fertilizers Know Nutrients	0.474	1.085	0.000	7.000	737
# Fertilizers Nutrients Correct	0.197	0.527	0.000	4.000	737
# Fertilizers Heard of	5.423	2.173	1.000	10.000	737
Use Urea	0.434	0.496	0.000	1.000	737
Use DAP	0.590	0.492	0.000	1.000	737
Use MOP	0.016	0.127	0.000	1.000	737

Figure 6: Baseline Summary Statistics

3.3.3 *Language* measures

The substantive portion of the lab-in-the-field began after the baseline. One set of measures captures whether treated farmers benefit from being provided with the *language of macronutrients* as mediators of fertilizer action. Given that baseline nutrient literacy is very low (Table 6), we expect provision of this *language* to be genuinely new for most treated farmers.

In particular, we measured how farmers substitute among fertilizers, exploit arbitrage opportunities, and whether they treated macronutrients as "decision-sufficient". We draw on our conceptual framework (Section 2) to show how performance on these activities would improve from a *language* of macronutrients, and illustrate the welfare benefits of each activity. Because optimal performance on these tasks doesn't require the *structure* of nutrient response, we view these *language* elicitations as a "first stage" of generalizing with mechanisms.

Fertilizer Substitution. This activity elicits whether respondents can identify fertilizer substitutes—alternatives that provide similar agronomic effects. In detail, respondents are presented with a vignette in which a farmer typically applies one plastic bottle cap of Urea per hole but is unable to find Urea at his agrodealer. They are asked to select another fertilizer (or set of fertilizers) and amounts that provide, as closely as possible, the same nutrients as the farmer's typical Urea application. Since Urea provides a single

macronutrient (N) in a known concentration, an appropriate substitute can be constructed by selecting another nitrogen-based fertilizer (such as CAN) in the appropriate quantity.

In terms of our framework, this presents an input x and asks respondents to choose x' such that $f(x') \approx f(x)$. We test whether respondents use the language of macronutrients to select $z(x') \approx z(x)$, i.e., one with approximately equal macronutrient content.

The ability to appropriately substitute fertilizers is important given the high frequency of stockouts in Sub-Saharan Africa (Vos et al., 2025). In our sample, 26% of respondents report at least one stockout of a preferred fertilizer in the past 12 months. Without this ability, farmers may choose inappropriate substitutes or forgo fertilizer application entirely.

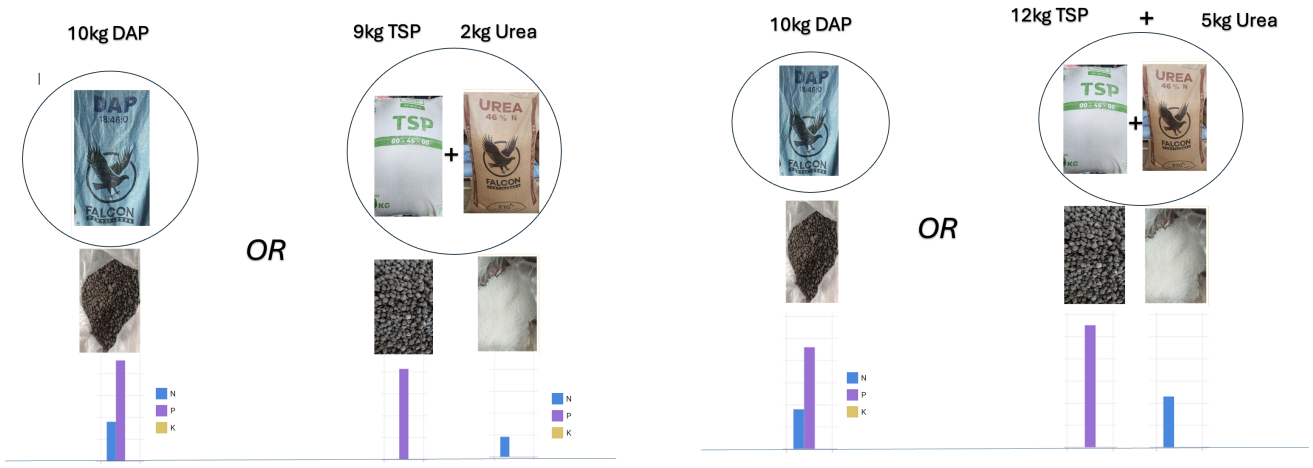
Fertilizer Arbitrage This activity elicits whether respondents exploit arbitrage opportunities by acquiring the same nutrients at lower cost. In detail, respondents make a series of binary choices between a 10kg bag of DAP and combinations of Urea and TSP. These fertilizers are selected because DAP provides both nitrogen (N) and phosphorus (P), while Urea and TSP are single-nutrient fertilizers providing N and P, respectively. Across choices, the quantities of Urea and TSP vary, starting from 9kg of Urea and 2kg of TSP and ending at 12kg of Urea and 5kg of TSP. In two elicitations, the DAP option is macronutrient dominant, and in two elicitations, the Urea–TSP combination is macronutrient dominant. We test whether respondents choose the option that provides strictly greater macronutrient content. The respondent-facing visual aid for this activity is shown in Figure 7.

In our framework, each choice presents inputs x and x' with macronutrient vectors $z(x)$ and $z(x')$. We test whether $z(x) > z(x') \Rightarrow x \succ x'$, i.e., whether respondents select the alternative that strictly dominates in macronutrient provision.

The ability to exploit nutrient arbitrage is important in settings with rapidly fluctuating fertilizer prices. In our sample, 52% of respondents report that the price of a fertilizer they wished to purchase was much higher than expected in the past 12 months. The ability to identify lower-cost alternatives providing equivalent nutrients can save farmers money and protect them against input market shocks (Vos et al., 2025).

Macronutrient Decision-Sufficiency This activity elicits whether respondents "dimension-reduce" fertilizer choices to their macronutrient content. If so, a respondent's predicted efficacy of an application of macronutrients should be invariant to the fertilizer recipe used to instantiate it.

Respondents are presented with two macronutrient-equivalent fertilizer plans: the first applies 10g of NPK 25-5-5 at transplanting; the second applies 3g of NPK 17-17-17 and 4.3g of Urea, also at transplanting. Enumerators read the plans aloud, show them in a table, and display the quantities on a common



(a) DAP bag is nutrient-dominant

(b) TSP-Urea bag is nutrient-dominant

Figure 7: Examples of respondent-facing visual aids for the Fertilizer Arbitrage activity. The enumerator presents the quantities of each "bag" of fertilizer, their appearance (bags and granules), and their macronutrient content (without mentioning which bag is nutrient-dominant).

scale using the palm of a hand. Respondents are also shown the implied macronutrients of each plan, which happen to be exactly equal. Examples of the respondent-facing visual aids for this activity are shown in Figure 8.

Farmers are then asked to predict the yield for each plan under the conditions of the demonstration plot. They report their yield predictions for each plan separately, and are subsequently asked to predict the difference in yields between the two plans. We elicit both individual yield predictions and the difference to increase precision and verify comprehension. We test whether farmers report small differences between the predicted yield of the two plans, and in particular whether treated farmers predict smaller yield differences between the two plans than control farmers.

In our framework, respondents are shown two inputs x and x' along with the fertilizers' macronutrient vectors $z(x)$ and $z(x')$, and asked to predict the corresponding outputs $f(x)$ and $f(x')$. Because the two plans provide identical macronutrients, i.e. $z(x) = z(x')$, the conceptual framework predicts similar yields, $f(x) \approx f(x')$. We assess whether respondents infer this equivalence by examining whether predicted yield differences $|f(x) - f(x')|$ are smaller in the treated group.

While the substitution and arbitrage activities test whether (treated) respondents attend to macronutrients as decision-*relevant*, it is a further step to see whether they treat them as decision-*sufficient* by dimension reducing to macronutrient equivalence classes. This would enable more efficient information aggregation across nominally distinct input bundles, improving inference from both one's own and others' outcomes.



Figure 8: Respondent-facing visual aid for the Decision-Sufficiency exercise. For each recipe (10g NPK 25-5-5 on the left; 3g NPK 17-17-17 + 4.3g Urea on the right), the enumerator presents quantities shown relative to the palm of a hand, fertilizer bag appearance, and macronutrient content (without mentioning that the two recipes are macronutrient-equivalent).

3.3.4 *Structure* measures

Our next set of measures captures whether treated farmers better leverage the *structure of nutrient response*, drawing again on our conceptual framework in Section 2. Many fertilizer choice problems—including targeting optimal fertilizer combinations, experimenting profitably, and avoiding poor choices—can benefit from this structural understanding.

Interactive Adaptive Experimentation Game Our most important *structure* measure is incentivized and mimics farmers' routine decision problem in a lab-in-the-field setting: choosing fertilizers for a growing season and iterating on those choices in subsequent seasons based on performance feedback. We elicited these choices on a custom digital application. The application placed respondents in a virtually unconstrained, realistic sandbox environment. Conceptually, they faced a finite-horizon multi-armed bandit problem, allowing us to generate granular inferences on both overall performance and how farmers navigate the exploration–exploitation trade-off.

In each of six "growing season" rounds, farmers could choose virtually unconstrained fertilizer recipes (including mixes) under common conditions (fixed complementary inputs and assigned soil context), track costs, and—if in the treatment group—view nutrient profiles. After finalizing a recipe for the round, they received instantaneous and incentivized feedback on simulated yield and profits, with earnings equal to 1% of the seasonal simulated profit on a quarter acre. For subsequent rounds, farmers could either copy their previous recipe or design a new one from scratch. The game ended after six rounds, with a die

roll determining which round was paid out: the maximum reward was economically meaningful at 10,000 UGX ($\sim$ 3 USD), roughly equivalent to a day's wage.

Screenshots are provided in Figure 9, and the app can be explored in "practice mode" at https://agriwork-82b57.web.app. A short "dummy" round was played first to familiarize farmers with gameplay. While the module was designed to be played "over the shoulder" with the farmer directing an enumerator, many farmers took full control of the device themselves, attesting to its intuitive design.[17]

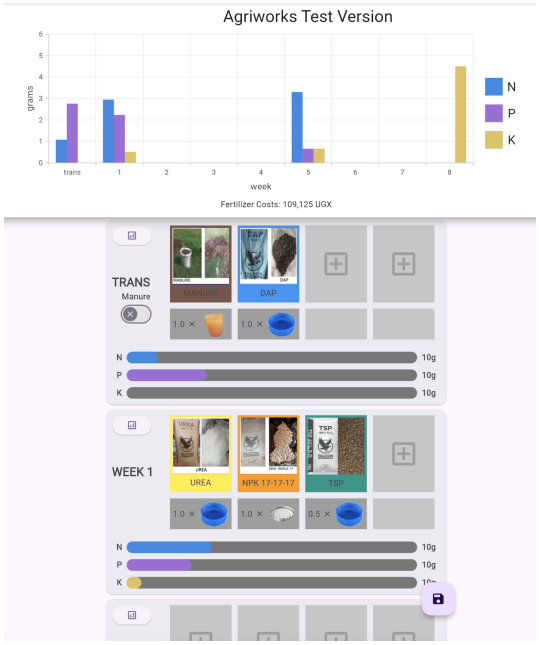
The experimentation environment is a structured finite-horizon bandit problem, where each fertilizer recipe constitutes a bandit arm. The agronomist's priors informed the expected reward for each arm, with independent Gaussian noise added to make the problem a bona fide stochastic bandit. Because the reward structure is linked across bandit arms (fertilizer recipes), the bandit is structured; better prior knowledge of the underlying structure (i.e. the agronomic mechanisms) should improve performance (Lattimore and Szepesvári, 2020). Due to technical issues, we were only able to deploy this module after the first four tranches, so the sample size for this activity is $N = 408$.

Predictions of Prespecified Recipes In a complementary incentivized measure, respondents were asked to predict the yields of several pre-specified fertilizer recipes with the understanding that their predictions would be compared to an agronomist benchmark. Each recipe involved a single fertilizer applied at one or two times, again in common complementary input conditions in the assigned soil context. Predictions were elicited before the demonstration (priors) and re-elicited after (posteriors). At the end of the survey, one prediction was randomly selected, and farmers received a 5,000 UGX ($\approx$ 1.4 USD) reward if their prediction was within 10% of the agronomist's benchmark.

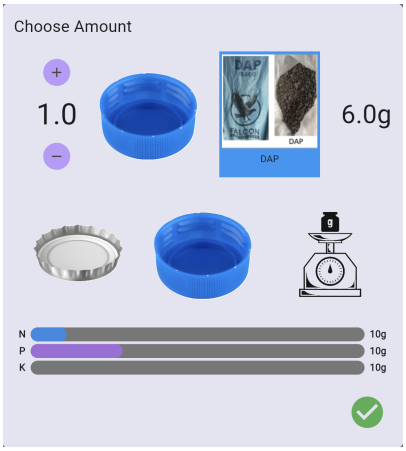
Recipes were presented in pairs that applied the same fertilizer type and quantity but varied the timing of application. Within each pair, one recipe aligned with the mechanism of nutrient response emphasized in the training, while the other did not. For example, the mechanistic explanation for phosphorus action stresses that plants can recycle phosphorus, so earlier application provides more benefit. Accordingly, applying the phosphorus-based fertilizer TSP at transplanting has higher marginal returns than applying the same quantity in week six. All the pairs of fertilizer recipes are shown below in Table 1.

We are interested not in the absolute accuracy of the yield prediction itself, but rather the directional accuracy between the yield predictions within each pair. Specifically, we hypothesize that this difference should lean more heavily in favor of the mechanism-aligned recipe for mechanistically trained farmers. In terms of our conceptual framework, we test whether for two input combinations (same quantity but different timing choices) x and x', where $g(z(x)) > g(z(x'))$ (i.e., where the structure function estimates a higher output for timing x than x'), the agent believes that $f(x) > f(x')$: the output of x is greater

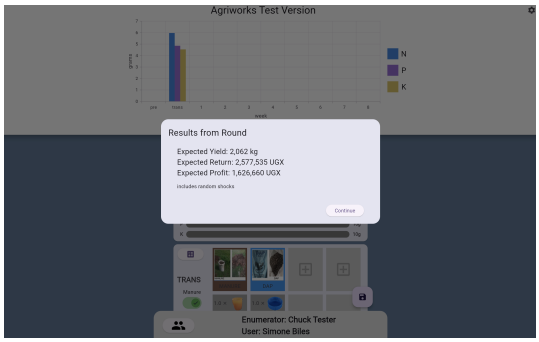
[17] We thank ProsperPlay (www.prosperplay.org) for developing and co-creating a wonderful tool.



(a) 10-week Recipe View with Nutrient Distribution



(b) Fertilizer Selection and Amount



(c) Recipe Yield, Revenue, and Profit

Figure 9: Adaptive Experimentation Digital Tool Interface

than the output of x'.

This exercise complements the adaptive experimentation task in two ways. First, it allows us to isolate belief updating, since we elicit beliefs before and after the demonstration. Second, because choices in the adaptive experimentation task reflect a combination of knowledge, risk preferences, and endogenous updating, these pre-specified comparisons provide a more direct measure of whether farmers understand mechanistic alignment across fertilizer timing.

Pair	Recipe	Fertilizer	Unit	Time	Aligns with mechanism?
1	1	TSP	2 plastic caps	transplanting	Yes
1	2	TSP	2 plastic caps	week 6	No
2	3	MOP	2 plastic caps	transplanting	No
2	4	MOP	2 plastic caps	week 6	Yes
3	5	Urea	4 plastic caps	transplanting	No
3	6	Urea	2 plastic caps	transplanting	Yes
		Urea	2 plastic caps	week 5	

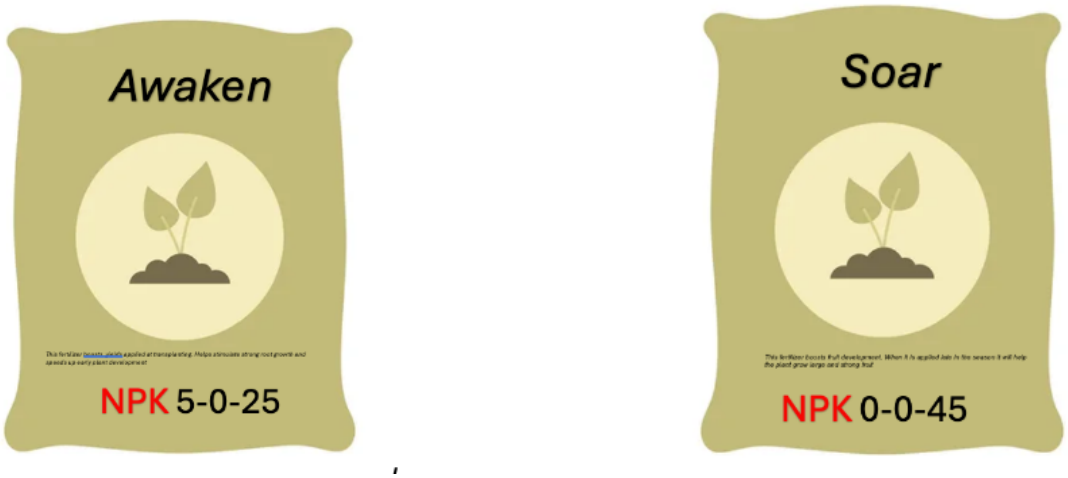
Table 1: Prediction Problem Fertilizer Recipes

Marketing Vignettes This activity elicits how farmers assess marketing claims about fertilizers that either cohere with, or contradict, their nutrient content. Respondents are presented with two marketing vignettes for fictitious fertilizers. Each vignette depicts a bag of fertilizer that includes a product name, logo, macronutrient (NPK) content, and a marketing claim describing expected effects. Respondents are asked to rate the credibility of each marketing claim.

The first vignette describes a fertilizer named "Awaken," with NPK number 5-0-25 and the claim: "This fertilizer helps boost yields when applied at transplanting. It helps stimulate strong root growth and speeds up early plant development." The second vignette describes a fertilizer named "Soar," with NPK number 0-0-45 and the claim: "When applied late in the season, it helps the plant grow large and strong fruit." The "Awaken" marketing claim mechanistically contradicts its revealed macronutrient content (it contains no phosphorus, the nutrient particularly helpful for root growth and early plant development). The "Soar" claim, on the other hand, mechanistically accords with its nutrient content (it contains potassium, which does help with fruiting). The respondent-facing visual aid associated with this activity is shown in Figure 10.

In our framework, respondents observe a proposed input x and nutrient content $z(x)$. They are given a claim about high returns $f(x)$ when applied at a certain time and can use their understanding of nutrient response $g(z(x))$ to evaluate that claim.

The ability to assess the credibility of marketing materials is important in this setting, where many products are marketed with hyperbolic names and descriptions. Understanding the mechanisms underlying these products can help farmers avoid being misled.



(a) Awaken: Inaccurate Marketing Claim

(b) Soar: Accurate Marketing Claim

Figure 10: Respondent-facing visual aid for the Marketing Vignettes. For each of the fictitious fertilizers (a) and (b), the enumerator walks through the name and the NPK content. Finally the enumerator points to the tiny text and reads aloud for the farmer the fertilizer's marketing claim.

Secondary *Structure* Outcomes We conduct other unincentivized activities to test for understanding of *structure* (without directly assessing knowledge), described in detail in Appendix A. In our *diagnosis and treatment* exercises, we test farmers' ability to diagnose nutrient deficiencies and recommend improvements. In the *recipe translation* exercise, we test understanding of directional changes of the demonstrated recipe in counterfactual soils, and additionally with and without manure to account for differential returns to manure across soils. In our *meta-mechanisms* elicitations, we test for uptake of simple implications of mechanistic understanding, guided by psychology findings that even when mechanistic recall decays, these corollaries may persist in memory and guide choice (Keil, 2006). Finally, in our *psychological attitudes* elicitations, we test whether mechanistic understanding is psychologically empowering, which may arise from improved understanding or reflect a purely behavioral response to the training.

3.3.5 WTP for Novel Fertilizer

We elicited willingness to pay (WTP) for a potassium fertilizer (MOP) included in the demonstrated recipe. While available in Uganda, this fertilizer was rarely found in local markets in the study region, and only 1.6% of attending farmers reported ever having used it. WTP for MOP thus measures demand for a novel fertilizer—a new "technology" within the recipe. We used an incentivized multiple-price

list elicitation. Farmers are offered the chance to purchase 2kg of MOP, enough to apply on a small test subplot. We present farmers with six price points, ranging from 0 UGX (free) to 15,000 UGX in increments of 3,000 UGX. The market price of 2kg of MOP is approximately 9,000 UGX, so this set of prices allows us to estimate demand for MOP in a range both above and below the market price. Following (Jack, McDermott and Sautmann, 2022*a*), we randomize both the order in which options are presented (ascending vs. descending price) and the order of responses (accept or reject the offer).

Mechanistic explanations may increase WTP for MOP through both improved knowledge and purely behavioral channels (such as a taste for flashy mechanisms or a higher reputational assessment of the extension officer). Knowledge-based increases may arise because farmers recognize MOP as a potassium fertilizer associated with improved yields. These gains may derive from either *language* or *structure*: for example, noticing that MOP has a high potassium concentration relative to other market fertilizers (*language*), or understanding that potassium supports fruit filling and late-stage growth (*structure*). However, the provision of the *language* may also introduce a countervailing force. If farmers recognize that potassium can be obtained from alternative fertilizers, they may view MOP as substitutable and reduce WTP. That said, other potassium-containing fertilizers typically have substantially lower potassium concentrations, and the limited availability of MOP in this region appeared to the researchers to reflect a market inefficiency.

3.4 Endline data

We conducted follow-up surveys in two agricultural seasons after the demonstrations to assess real-world impacts. We successfully re-interviewed 92% of demonstration attendees, with no evidence of differential attrition; 75% of these cultivated tomato plots in at least one of the two seasons. Our primary outcomes focus on detailed inputs and outputs on respondents' tomato plots. Our secondary outcomes capture broader cognitive and behavioral responses, including knowledge, belief updating, and social interactions.

3.4.1 Primary outcomes

On up to two of their tomato plots (restricting to the top two by size if more than two), we administer a detailed module eliciting farmers' input choices, input costs, yields, revenues, and any subplot experimentation. The most granular elicitation concerns fertilizer decisions, where we collect the full fertilizer schedule, including types, quantities, and timing of application. A more limited elicitation is administered for non-tomato plots, capturing crop choice and cultivation status but not detailed input use or outputs. All outcomes are based on self-reports collected near the end of each growing season.

3.4.2 Secondary outcomes

We also capture belief updating about fertilizers, interactions with peers and the extension officer, and whether these conversations pertained to fertilizer use. In addition, because our intervention provides mechanistic explanations in an environment where such information is typically scarce, we elicit demand for mechanistic information about another agricultural technology (hybrid seeds). Specifically, we implement an incentivized willingness-to-pay exercise to watch a brief video that explains hybrid vigor (heterosis) through basic Mendelian genetics. This allows us to assess whether exposure to mechanistic explanations engenders a broader demand for mechanistic information and curiosity beyond the specific technology studied.

4 Results

4.1 Lab-in-the-field Results

We organize the results into two broad parts based on our conceptual framework. *Language* tests examine whether farmers treat macronutrients as decision-relevant features of fertilizer impacts, while *structure* tests examine whether they apply this knowledge to reason about timing, interactions, and adaptation across contexts.

Because our design is a location-stratified RCT at the farmer level, the basic estimation strategy is straightforward:

$$y_{ist} = \alpha + \beta \,\text{Treat}_i + \gamma_s + \delta_t + \theta' \mathbf{Z}_i + \varepsilon_i, \tag{1}$$

y_{ist} is the outcome for farmer i, in location stratum s, and who attended the training in tranche t. γ_s and δ_t are location-stratum and date-tranche fixed effects, respectively. We include date-tranche fixed effects to account for any potential non-random variation in farmer responses due to date of attendance (each tranche consists of two consecutive dates). In all specifications we also include controls for baseline farmer characteristics selected through double-post LASSO $\mathbf{Z}_i$ following Belloni, Chernozhukov and Hansen (2014). We estimate heteroskedasticity-robust standard errors but do not cluster errors above the farmer level, since randomization was assigned at the individual level.

We sometimes explore an interacted variant

$$y_{ist} = \alpha + \beta_1 \,\text{Treat}_i + \beta_2 \mathbf{X}_i + \beta_3 \text{Treat}_i \mathbf{X}_i + \gamma_s + \delta_t + \theta' \mathbf{Z}_i + \varepsilon_i, \tag{2}$$

where $\mathbf{X}_i$ is some farmer-level pre-demonstration quantity of interest (e.g. baseline covariates or prior beliefs).

4.1.1 *Language* Results

With reference to the elicitations described in Section 3.3.3, we document whether treated farmers better substitute for missing fertilizers, exploit nutrient arbitrage opportunities, and treat macronutrients as decision-sufficient.

Table 2: Fertilizer Substitution Problem

	(1) Index	(2) N-distance	(3) NPK-distance	(4) Choose CAN	(5) Choose Low-N Fertilizer
Treat	0.083	-0.135	0.047	0.044	-0.083
	(0.042)	(0.074)	(0.189)	(0.027)	(0.033)
	[0.045]	[0.066]	[0.804]	[0.111]	[0.012]
Observations	797	797	797	797	797
Control Mean	-0.0977	1.513	3.318	0.173	0.391
Control SD	0.604	0.897	2.851	0.378	0.489

Fertilizer Substitution Table 2 presents the results on how appropriately respondents substitute for an unavailable cap of Urea. Column (2) estimates the effect of the treatment on the distance between the quantity of nitrogen in the farmer's selected substitute and the target amount of nitrogen (in units of grams per plant). The treatment decreases this distance by 9% relative to the control mean, an effect size of 0.15 SD ($p = 0.066$). Treated farmers were better able to select substitutes that provided similar amounts of the nutrient provided in the target fertilizer. Column (3) displays results for Euclidean distance between the nutrients in the substituted fertilizer and the target fertilizer. The coefficient here is slightly *positive* but insignificant. Although treated farmers chose substitutes that better matched the nitrogen content of the target fertilizer, these substitutes on average provided extraneous nutrients, attenuating the overall match in NPK space. These results are consistent with Column (4), which displays results for a binary indicator of whether the farmer chose CAN, the only other nitrogen-only fertilizer available. The effect here is positive but insignificant. This suggests that the result on more accurate substitution of nitrogen may be partially driven by choosing the nitrogen-only alternative, but not exclusively. Column (5) provides further evidence. The outcome here is a dummy variable for whether the respondent chose one of the two fertilizer options that *primarily provide a nutrient other than nitrogen*, the target nutrient (DAP and MOP, which provide primarily P and K, respectively). The treatment reduces the probability of choosing a poor substitute fertilizer by 8.3 percentage points ($p = 0.012$).

Overall, the results indicate that mechanistic explanations improved substitution, though farmers appear to focus primarily on the salient nutrient in the unavailable fertilizer (nitrogen) rather than jointly

optimizing across nutrients.

Fertilizer Arbitrage Figure 11 shows the distribution of the number of correct responses for the treatment (mechanisms) and control (signal-only) groups, where "correct" refers to selecting the nutrient-dominant option among a pair of alternatives. Treated farmers are significantly more likely to choose the option that provides more nutrients in all four choices. Combining the responses to these four questions into a standardized index, we find that treatment increases performance on the standardized index by 0.11 SDs ($p < 0.01$).

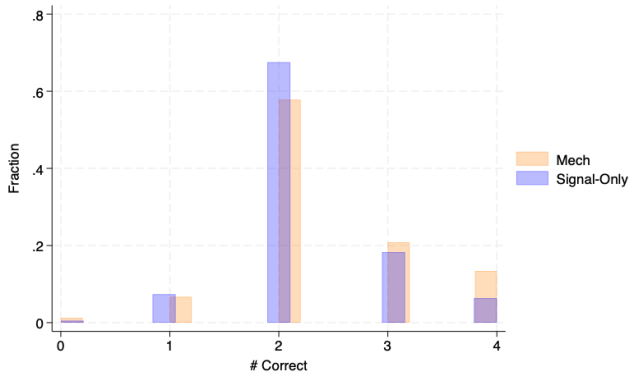


Figure 11: # Arbitrage Questions Correct

Macronutrient Decision-Sufficiency Table 3 presents results on respondents' predictions of yield differences between macronutrient-equivalent recipes. Columns (2) and (3) report differences in predicted yields calculated in two distinct ways. In Column (2), the difference is computed in the backend after respondents report yields for each recipe sequentially, while Column (3) uses respondents' direct elicitation of the difference. Column (1) reports an index constructed following (Kling, Liebman and Katz, 2007). The table shows no significant treatment effect on the difference in predictions, despite large control mean differences (296 kg and 193 kg in Columns (2) and (3) respectively). A kernel density plot in Figure 15 (Appendix A) indicates that this null result does not mask meaningful distributional impacts.

This result suggests that treated farmers were not more likely to treat macronutrients as decision-sufficient. Especially in light of earlier findings that the treatment increases the decision relevance of macronutrients, this raises an important puzzle for future research: do farmers resist dimension reduction altogether, or only for these particular recipes? For these particular recipes, the large reported differences (as reflected in the control means) represent a clear failure of agronomic understanding, since these recipes contain no micronutrients and are often formulated from the same base fertilizers (e.g., Urea, DAP, MOP), and can therefore be *literally identical*. Resisting dimension reduction in general, however, is more nuanced. Respondents may have failed to deduce the dimension reduction, which would constitute an

objective failure, or the behavior may reflect the extent to which they view the mechanistic representation as incomplete—as all representations are.[18] In the case of synthetic fertilizers, dimension reduction is largely justified because products are explicitly formulated to supply particular nutrients, while differences in how these nutrients are released across formulations are typically minor. This suggests that the dimension-reducing deduction may need to be made explicit in the intervention.

Table 3: Difference in Yield Predictions for Nutrient-Equivalent Fertilizers

	(1) Diff Index	(2) Sequential	(3) Direct
Treat	0.014	-16.588	9.545
	(0.055)	(23.855)	(16.859)
	[0.801]	[0.487]	[0.571]
Observations	797	797	797
Control Mean	-4.29e-09	295.8	193.5
Control SD	0.831	378.3	243.2

Findings. Treated farmers perform significantly better on the Substitution and Arbitrage indices. They are more likely to identify equivalent substitutes (e.g., CAN for Urea) and to make choices that exploit nutrient-based arbitrage opportunities. However, they do not collapse fertilizer choices fully into macronutrient-equivalence classes. This suggests that farmers treat macronutrients as *partially sufficient*, but do not adopt the full abstraction.

4.1.2 Structure Results: Adaptive Experimentation Task

Here we document whether treated farmers have improved performance on the naturalistic fertilizer application activity described in Section 3.3.4, the aspects on which they improve, and the distributional impacts.

We modify the estimating equation slightly for our analysis of the adaptive experimentation task, to account for the round-based structure of the data. As before, we estimate simple treatment-effects regressions but now the unit of observation is the respondent-round level, where rounds are indexed by r.

$$y_{irst} = \alpha + \beta \, \text{Treat}_i + \gamma_s + \delta_t + \zeta_r + \theta' \mathbf{Z}_i + \varepsilon_i, \tag{3}$$

Since farmers will learn from feedback in previous rounds, we also include round fixed-effects (ζ_r) to

[18] (Levi-Strauss, 1966) develops a related intuition in the notion of "bricolage," where a form of conceptual knowledge is applied pragmatically within limited domains, reflecting epistemic humility about extrapolating beyond its band of applicability.

absorb this variation and improve power. We also include the randomly assigned soil type in the vector of covariates Z_i. Since we observe multiple observations for each respondent, we cluster standard errors at the respondent level.

Table 4: Profits in Adaptive Experimentation Task

	(1) Profits	(2) Profits
Treat	20.731	29.261
	(10.406)	(17.734)
	[0.046]	[0.099]
Grey Soil	-333.061	-327.039
	(10.373)	(15.904)
	[0.000]	[0.000]
Grey Soil × Treat		-14.063
		(21.830)
		[0.519]
Observations	2445	2445
Control Mean	240.0	240.0
Control SD	233.4	233.4
Het. Var Mean		0.493
Mech + Mech × Het. Var = 0		0.215

Profits Table 4 presents results on average treatment effects on profits (in 1000s of UGX). The result in Column (1) shows that treated farmers earned an additional profit of 20,731 UGX compared to control farmers. This effect is statistically significant ($p = 0.046$) and reflects a 9% increase in profits over the control mean. This effect does not appear to be concentrated among farmers who play the game on either the black or grey soil; although average profits are significantly higher on the black soil, the interaction term between the treatment and the grey soil is negative and insignificant.

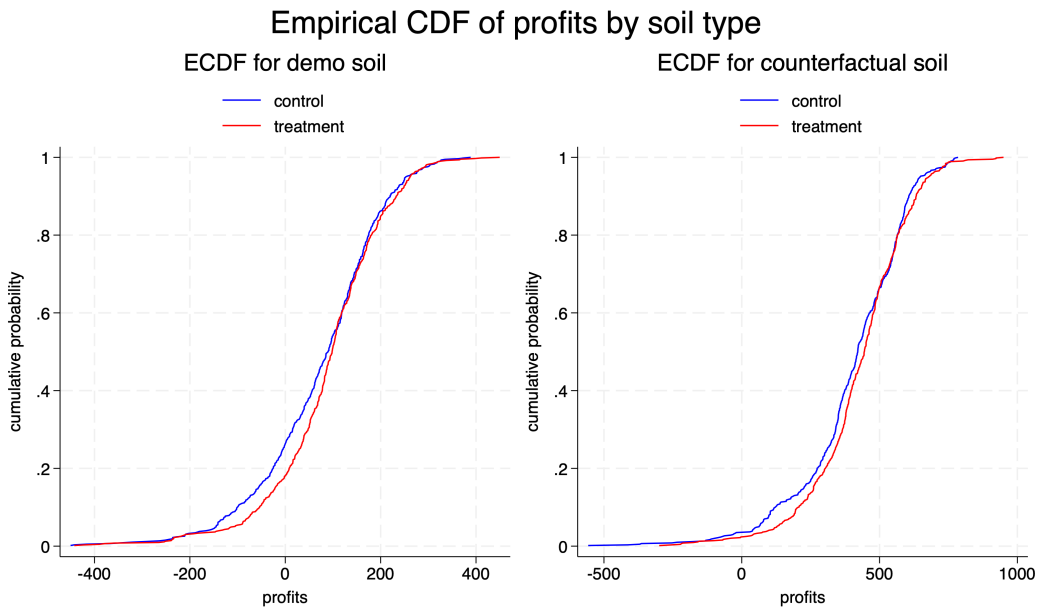


Figure 12: Empirical CDF of profits per soil type

Turning to distributional impacts, we first examine how the distribution of profits shifts. Figure 12 plots empirical CDFs of treatment and control profits. Treatment profits appear stochastically dominant on both soils. The gains appear to come primarily from a reduction in losses and an increase in moderate profits, with a smaller increase in large profits.

A further question is how treatment effects vary across farmers. The object of interest is the conditional average treatment effect (CATE) as a function of baseline covariates, which is, however, high dimensional, rendering naive estimation approaches (such as subgroup interactions) prone to overfitting. To address this, we apply the data-driven method of (Chernozhukov et al., 2025), which estimates key projections of the CATE onto a machine-learning proxy. We consider flexible learners—random forest and elastic net—to construct this proxy, both of which carry a signal of CATE heterogeneity in a best linear predictor exercise.[19] We then compute grouped average treatment effects (GATES), which estimate the mean of the CATE across quintiles of the proxy predictor (Figure 13).

[19] In the notation of (Chernozhukov et al., 2025), we reject $\beta_2 = 0$ in the best linear predictor results in Appendix Table A5

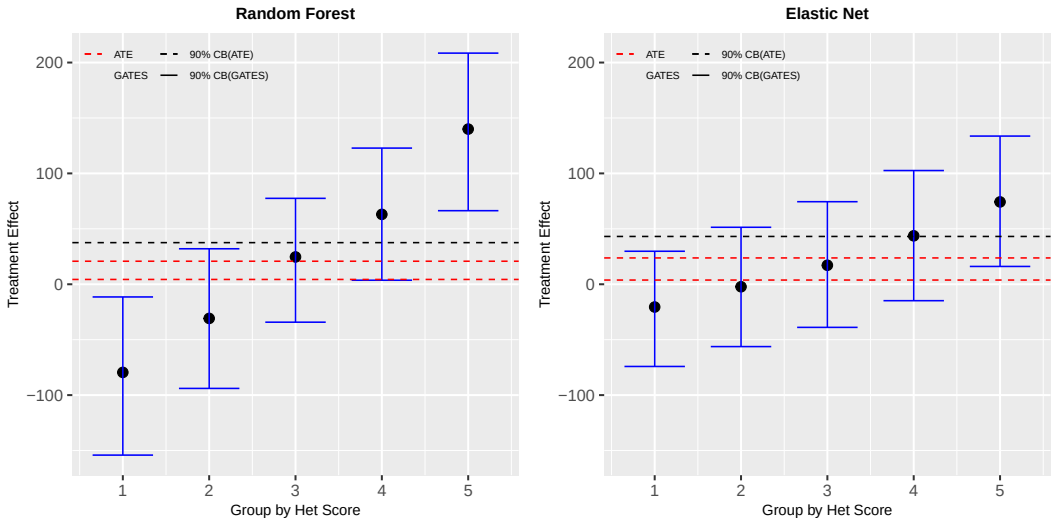


Figure 13: GATES (projection of the CATE onto quintiles of an ML proxy) following (Chernozhukov et al., 2025), with Random Forest (left) and Elastic Net (right). 200 simulations.

Several patterns emerge. First, the random forest proxy detects substantially more heterogeneity, consistent with the goodness-of-fit diagnostic of (Chernozhukov et al., 2025) reported in Appendix Table A6. We therefore focus on the left panel. Second, GATE estimates vary widely across quintiles, ranging from approximately −80 to +140 (thousands of UGX). The top two GATE quintiles are positive and statistically significant, and the highest quintile is substantially larger than the overall ATE (about seven times larger). Third, the lowest GATE quintile is negative and statistically significant, indicating a group of farmers for which the treatment *reduced* profits in this task.

To characterize these groups, we conduct classification analysis (CLAN) following (Chernozhukov et al., 2025), where we calculate baseline covariate means within each GATES group and highlight differences between top and bottom GATES that are statistically significant. The full set of statistically significant CLAN differences per the random forest ML proxy is shown in Appendix Table A7. Relative to the lowest quintile, the highest-GATE group was more educated, more experienced with advanced irrigation, had experimented more with fertilizers, and had greater familiarity with potassium fertilizers, suggesting that prior exposure may complement mechanistic training. By contrast, the lowest-GATE group had more experience with tomatoes and were less risk averse. These patterns may even suggest multiple potential mechanisms underlying heterogeneous effects—for example, more experienced tomato farmers may have been destabilized by new information, while lower risk aversion may reflect a willingness to "muck around" and try choices they may already suspect are suboptimal, in this relatively low-stakes setting. We view these interpretations as suggestive and leave further inquiry into potential "backfiring" effects of mechanistic explanations to future work.

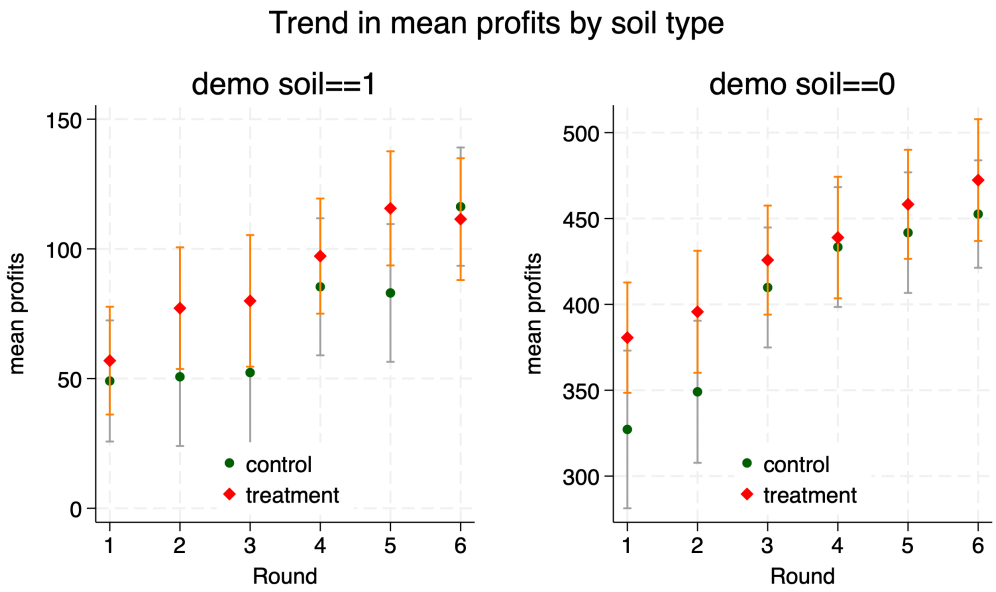


Figure 14: Trend in mean profits over the bandit rounds

Finally, we plot trends in farmers' profits over game rounds ("growing seasons"). Treatment profits are higher from the outset, and the control group partially narrows the gap over time. Since sustained experimentation toward an optimum is rarely observed in real-world smallholder farming settings, this convergence likely reflects the low-stakes nature of the game environment.[20] In that sense, the pattern may also underscore the pedagogical value of interactive digital games, which can induce exploration that would otherwise be costly in the field (Tjernström et al., 2021). At the same time, the fact that the control group eventually observes some high signals helps alleviate concerns that the lab-in-the-field itself generated differential endogenous exposure beyond the demonstration that could otherwise influence endline outcomes.

Yields and Costs Table 5 provides further evidence on performance in the bandit problem, showing impacts of the treatment on yields and costs in each round. Columns (1)-(2) present results on yields and Columns (3)-(4) presents results on costs. The treatment leads to a moderate increase in yields, although the p-value is slightly above the threshold for significance at the 10% level. Turning to costs, we find a small reduction in cost per round in Column (3), although this result is also slightly above the threshold for significance at the 10% level. Regardless of the imprecision, the direction of these estimates on yields and costs suggests that treated farmers are not "juicing the inputs". In fact treated farmers apply weakly less fertilizer, without incurring a yield penalty.

[20]Indeed, most respondents had farmed and used fertilizers for years, yet baseline fertilizer practices often fail basic agronomic benchmarks.

Table 5: Yields and Costs in Adaptive Experimentation Task

	(1) Yield	(2) Yield	(3) Costs	(4) Costs
Treat	19.333	25.644	-4.542	2.345
	(13.761)	(22.251)	(3.100)	(4.671)
	[0.160]	[0.249]	[0.143]	[0.616]
Grey Soil	-409.800	-403.294		5.291
	(13.865)	(19.333)		(4.320)
	[0.000]	[0.000]		[0.221]
Grey Soil × Treat		-12.757		-13.876
		(27.715)		(6.211)
		[0.645]		[0.025]
Observations	2445	2445	2445	2445
Control Mean	869.2	869.2	455.4	455.4
Control SD	290.3	290.3	45.37	45.37
Het. Var Mean		0.493		0.493
Mech + Het. Var × Mech		0.429		0.00445

Nutrient Applications The evidence presented so far on profits, yields, and costs suggests that treated farmers are choosing fertilizer applications that align more closely with the agronomy of plant growth, boosting profits through both increasing yields and reducing costs. We can probe this more directly by examining alignment with agronomy-based decision rules for the timing of nutrient application for each macronutrient. For nitrogen, the rule is to apply nitrogen during the vegetative stage of plant growth, which is important for leaf and stem growth as well as for splitting applications across multiple periods (since most farmers already apply nitrogen at planting). For phosphorus, the rule is to apply it at planting, since the nutrient can be recycled many times by the young plant. For potassium, the rule is application during the fruiting stage of growth, since potassium is essential for fruit set and quality. For each nutrient, we calculate the percentage of the total nutrient application that was applied during the target time period. We also construct a standardized index of these percentages to adjust for multiple hypotheses and to increase power.

These results are presented in Table 6. The impacts for each nutrient application rule are positive. The impacts for % of nitrogen applied during the vegetative growth stage (Column 1) and % of phosphorus applied during the planting stage (Column 2) are positive but imprecisely estimated. For % of potassium applied at fruiting, this result is large and statistically significant at the 5% level (Column 3). The result implies that treated farmers applied 6 percentage points more of their total potassium application at the fruiting stage, an increase of 13% over the control mean. Column (4) shows impacts of the treatment on the standardized index of these application percentages. The impact is positive and statistically significant at

Table 6: Mechanism-Aligned Application in Adaptive Experimentation Task

	(1) % N Vegetative	(2) % P Planting	(3) % K Fruiting	(4) Index
Treat	1.330	1.231	5.755	0.101
	(1.329)	(1.719)	(2.680)	(0.038)
	[0.317]	[0.474]	[0.032]	[0.009]
Observations	2444	2439	2403	2444
Control Mean	10.91	81.47	43.26	0.0146
Control SD	16.27	23.06	34.69	0.486

the 1% level ($p < 0.01$). Overall, these results suggest that farmers in the mechanistic treatment adjusted their fertilizer applications to more closely align with the mechanisms of tomato growth, helping them earn higher profits.

4.1.3 Direct *Structure* Results

While the adaptive experimentation task reflects a compound activity in which knowledge of structure interacts with preferences (e.g. risk preferences), we next turn to direct structure measures in Section 3.3.4, which more narrowly isolate reasoning about nutrient response.

Table 7: Structural Prediction Problems

	(1) Index	(2) Index
Treat	0.057	0.050
	(0.044)	(0.049)
	[0.195]	[0.306]
High Priors × Mech		0.136
		(0.112)
		[0.225]
High Priors		0.143
		(0.185)
		[0.438]
Observations	694	694
Control Mean	0.00648	0.00648
Control SD	0.597	0.597
Mean (unstandardized)	18.70	18.70
Mean(Het. Var)		0.332
Mech + Het. Var × Mech = 0		0.0654

Prediction of Prespecified Recipes Table 7 presents results on how respondents predict yields of prespecified recipes. The dependent variable in each column is the index of percentage yield differences, constructed as follows. We first compute the percent difference between yield predictions within each pair of recipes, rescaling so that positive values indicate a higher yield prediction for the mechanism-aligned recipe. We then combine the differences for each pair into a standardized index following the procedure of Kling, Liebman and Katz (2007).

The estimate in Column (1) indicates a small but insignificant positive impact of the treatment.Column (2) examines heterogeneity by baseline beliefs, in particular we interact the treatment indicator with a dummy for whether the standardized index of prior beliefs (constructed in exactly the same way as the outcome index, but using the beliefs elicited at baseline) is above zero (the mean). The interaction effect is positive and the sum of the coefficients of the treatment variable and the interaction term is significant at the 10% level ($p = 0.0654$). This implies that the treatment increases the standardized index by 0.2 points for farmers with prior beliefs that favor the mechanism-aligned recipes. In other words, the mechanistic treatment seems to cause farmers to update beliefs positively about the returns to mechanistically-aligned fertilizer applications, but this effect is concentrated only among farmers whose prior beliefs are relatively aligned with such mechanisms to begin with.

Although the mechanisms treatment reinforces beliefs about returns to mechanism-informed applications only among farmers who have prior beliefs about these returns, it still improves the overall accuracy of beliefs. For control farmers with above average prior beliefs, the mean posterior belief about the percentage return of the mechanism-aligned recipe over the non-mechanism-aligned recipe in each pair is 20.9%. However, according to the predictions of our agronomic model, the "true" average return is 41%. This suggests that the mechanistic treatment shifted beliefs about returns closer to the agronomically-sound prediction, reducing overall error in beliefs about returns.

The apparent discrepancy between effects on agronomic understanding revealed on this prespecified prediction problem, where it only appears for those with high priors, and the adaptive experimentation task, where it shows up as a positive average treatment effect, a first-order stochastically dominant CDF of profits, and two positive GATES, is striking. It is consistent, however, with recent evidence that people can do much better on naturalistic problems than on stylized abstract problems (Banerjee et al., 2025). The prediction exercise typically isolates a single fertilizer at a fixed time, a narrowly defined sub-problem closer to "math in school." By contrast, the adaptive experimentation challenge requires assembling a full fertilizer strategy aimed at profits, much closer to "math in the market." This contrast highlights the value of our digital tool as an unconstrained choice environment for observing how mechanistic understanding guides integrative decision-making.

Table 8: Marketing Vignettes

	(1) Index	(2) Misleading	(3) Accurate
Treat	-0.004	-0.017	0.012
	(0.044)	(0.031)	(0.027)
	[0.936]	[0.582]	[0.664]
Observations	797	797	797
Control Mean	9.08e-10	0.299	0.180
Control SD	0.624	0.459	0.385
Mean (unstandardized)	0.240		

Marketing Vignettes Table 8 presents results on how farmers assess accurate or inaccurate marketing claims given macronutrient content. The outcome variable in Columns (2) and (3) is an indicator for whether the respondent assessed the marketing claim correctly. Column (2) presents results for the first vignette, whose marketing claim is misleading, and Column (3) presents results for the second vignette, whose claim is accurate. Overall performance on evaluating marketing claims is poor. In the control group, fewer than a third of respondents correctly identified the first vignette as misleading, and fewer than a fifth of respondents correctly identified the second vignette as accurate. Despite this low baseline, the treatment does not improve performance. Estimated treatment effects in both columns (2) and (3) are small and statistically insignificant. We also do not find evidence of significantly different impacts of the treatment for low- and high-baseline knowledge farmers (results not shown). These results may indicate that farmers had difficulty extracting relevant information on macronutrient content from the NPK numbers (e.g., 5-0-25), despite this being part of the mechanistic curriculum. However, it is also possible that farmers parsed the NPK information correctly but did not treat it as *decision-sufficient* (as suggested by the analysis of Table 3 in Section 4.1.1), leaving room for other "hidden" active ingredients that could rationalize misleading claims.

Taken together, these findings suggest that, despite the evidence on nutrient arbitrage, mechanistic explanations alone are unlikely to fully "arbitrage away" the profusion of branded fertilizers in the market. Extension officers may need to make this implication explicit—for example, emphasizing that branded formulations often don't offer much beyond the NPK they provide.

WTP for Novel Fertilizer Table 9 reports the average willingness to pay for 2kg of MOP, estimated through the random utility logit model of (Jack, McDermott and Sautmann, 2022*b*). The average WTP is 4,782 UGX (Column 1), slightly more than half the market price. The estimated treatment effect is positive but small and statistically insignificant. Exposure to mechanistic explanations of the recipe—which

Table 9: WTP for Novel Fertilizer

	(1) WTP	(2) WTP
Constant	4782.193	4458.529
	(464.167)	(628.550)
	[0.000]	[0.000]
Treat	80.773	631.386
	(342.578)	(644.561)
	[0.814]	[0.327]
Mover score × Treat		-5610.741
		(5677.084)
		[0.323]
Mover score		3283.820
		(3949.614)
		[0.406]
Observations	4782	

included the novel fertilizer—did not increase average demand.

In Column (2), we examine whether improved arbitrage ability may act as a countervailing force to increasing WTP. Specifically, we test for heterogeneity by the predicted treatment-induced increase in arbitrage ability. We don't interact directly with arbitrage ability, since this is a post-treatment endogenous variable. Instead, we estimate a "mover score" from baseline covariates that predicts treatment-induced increases in arbitrage ability—that is, propensity to be a "complier" of high arbitrage—using cross-validated logistic elastic net:

$$\text{Mover Score}(Z) = P(\text{high arbitrage} \mid T = 1, Z) - P(\text{high arbitrage} \mid T = 0, Z),$$

and interact this measure with treatment. The interaction effect estimate is negative, consistent with improved arbitrage ability dampening demand, but the estimate is statistically insignificant.

4.2 Secondary *Structure* Results

In Appendix A, we examine secondary structural reasoning measures. The treatment improves performance in the *diagnosis and treatment* exercise, consistent with reasoning agronomically about bottleneck nutrients. In the *recipe translation* task, we find suggestive evidence of improved extrapolation across soil contexts. Effects on *meta-mechanism* beliefs are positive but just above conventional significance thresholds ($p = 0.105$), though we observe a significant increase in emphasis on timing over quantity. We find null effects across measures of *psychological attitudes*, though strong indicators of experimenter

demand effects in this elicitation suggest that the elicitation may mask any underlying changes. Taken together, these results provide complementary evidence that mechanistic training improves key dimensions of structural reasoning beyond the adaptive experimentation task.

4.3 Endline

The endline results on farmers' real practices, yields, and gross profits (revenue minus costs) reflect more agronomically aligned fertilizer choices and higher yields. The estimation strategy is:

$$y_{ipwst} = \alpha + \beta \,\text{Treat}_i + \eta_w + \gamma_s + \delta_t + \zeta \mathbf{X}_i + \theta' \mathbf{Z}_i + \varepsilon_i, \tag{4}$$

Here y_{ipwst} is the outcome for farmer i on tomato plot p interviewed in wave (growing season) w, in location stratum s, who attended the training in tranche t.[21] We include wave, strata, and tranche fixed effects, and standard errors are clustered at the respondent level. We include double-post LASSO-selected baseline controls $\mathbf{Z}_i$ following (Belloni, Chernozhukov and Hansen, 2014), and also a small set of baseline controls $\mathbf{X}_i$ that we expect *a priori* to be prognostic of the outcome. In particular, for fertilizer-choice outcomes, we control for the farmer's self-reported soil archetype (grey, brown, or black). For yield and profit outcomes, we control for baseline average yields per acre.

Fertilizer Input Choices Table 10 reports key aspects of farmers' macronutrient allocation at endline. These nutrient quantities are calculated indirectly from farmers' fertilizer applications.

Table 10: Fertilizer Choices at Endline

	(1) % N vegetative	(2) % P planting	(3) % K fruiting	(4) Max N concentration
Treat	0.751	-2.797	4.177	-3.141
	(1.673)	(2.462)	(2.489)	(1.423)
	[0.653]	[0.256]	[0.093]	[0.027]
Observations	970	970	970	970
Control Mean	18.21	71.29	25.55	71.04
Control SD	24.85	33.30	35.73	20.48

The first three columns parallel the measures from Table 6 on the adaptive experimentation game: the relative emphasis on nitrogen, phosphorus, and potassium at their agronomically critical stages (vegetative, planting, and fruiting, respectively). Column (1) shows a small, imprecise increase in nitrogen emphasis

[21] We collected this detailed input and outcome data on up to two tomato plots per farmer per wave. In case the farmer had more than two tomato plots, we selected the largest two tomato plots.

at the vegetative stage (p=0.653). Column (2) shows a small, imprecise negative effect on phosphorus emphasis at planting (p=0.256), though the control mean is already 71%, reflecting the widespread use of DAP at the start of the planting schedule.[22] Column (3) shows that treated farmers apply 4.2pp (a 16% increase relative to the control mean) more of their potassium at the fruiting stage (p=0.093). A higher fruiting-stage potassium share is also a key mediator of the treatment group's higher simulated profits in the adaptive experimentation game (Table 6).

Column (4) turns to nitrogen balance. The outcome is the largest order statistic of the fractions of nitrogen applied at planting, vegetative, and fruiting stages; lower values indicate that nitrogen is spread more across stages rather than concentrated in one. Treated farmers spread their nitrogen applications more (p=0.027), consistent with the agronomic principle that nitrogen must be reapplied due to leaching risks. Together, these results show that treatment shifted fertilizer choices toward more agronomically aligned timing and balance—especially higher potassium at fruiting and a more distributed nitrogen profile.

These effects on potassium at fruiting and nitrogen balance are not driven by the reinforcement training between growing seasons. In Table A8, we estimate a specification that isolates the marginal effect of reinforcement over and above the treatment. The incremental effect of reinforcement is consistently indistinguishable from zero. For nitrogen splitting in particular, the treatment effect remains precisely estimated when the reinforced sample is excluded from the second wave (p <0.05).

Yields and Profits Table 11 reports downstream outcomes: yields per acre (metric tons per acre) and inverse-hyperbolic-sine–transformed gross profit (gross profit in UGX, calculated as revenue minus costs, without pricing out family labor). Relative to yields, profit is an inherently noisier construct, as it aggregates multiple input categories, each measured with error. We report specifications with and without controls for plausibly exogenous shocks (droughts, floods, disease, etc.).
Column (1) shows that treated farmers achieve 14% higher yields per acre (p=0.067). Adding shock controls in Column (2) slightly raises the estimated effect to 15% and improves precision (p=0.051). Columns (3)–(4) show positive but noisy effects on gross profit ($p > 0.5$).

These effects on yields are not driven by the reinforcement training. In Table A8, we estimate a specification that isolates the marginal effect of reinforcement over and above the treatment. The incremental effect of reinforcement is indistinguishable from zero, and the treatment effect remains precisely estimated when the reinforced sample is excluded from the second wave ($p < 0.10$).

[22] This is also consistent with our qualitative observations that farmers are ubiquitously told that "DAP is good for the roots."

Table 11: Yields and Profits at Endline

	(1) Yields per Acre	(2) Yields per Acre	(3) Gross Profit (ihs)	(4) Gross Profit (ihs)
Treat	0.882	0.935	0.517	0.476
	(0.481)	(0.478)	(0.839)	(0.809)
	[0.067]	[0.051]	[0.538]	[0.557]
Control Shocks	No	Yes	No	Yes
Observations	1000	1000	1002	1002
Control Mean	6.370	6.370	6.929	6.929
Control SD	6.209	6.209	12.22	12.22

Summary of Findings Mechanistic explanations induced more agronomically aligned fertilizer decisions at endline—particularly greater fruiting-stage potassium and a more distributed nitrogen schedule. These input changes coincide with 14–15% higher yields per acre, without a statistically detectable increase in gross profits.

5 Conclusion

In a fertilizer choice experiment, we find that farmers with low levels of formal education grasp and apply mechanistic explanations: they substitute fertilizers more accurately, exploit arbitrage opportunities, generalize from mechanisms to update their beliefs about the returns to fertilizers, and apply agronomically sound recipes to increase yields while lowering costs, achieving higher profits in an incentivized task. In follow-up data from farmers' own plots, we also observe shifts toward more agronomically aligned fertilizer practices and higher yields. Our findings suggest that the impacts of information interventions in complex domains are more powerful when combined with explanations of the mechanisms and principles that underlie their recommendations. Economic agents do not benefit only from the *what*, but also from the *why*.

Because our control group receives the same recipe demonstration but without the mechanistic training, we can collect systematic evidence on what mechanistic explanations add above a reduced-form signal. Given that we see farmers generalize—including across soil contexts, seemingly mitigating thorny issues of learning in heterogeneous environments (Alidaee, 2023; Chakraborty, 2022; Munshi, 2004; Tjernström, 2017)—this raises broader questions. To what extent do mechanistic explanations substitute for data, and to what extent do they complement it? Understanding these questions could inform the design of information campaigns across many domains – including health, agriculture, entrepreneurship, and education. In contexts where providing tailored or localized information is expensive, such as precision agriculture

and personalized health recommendations, mechanistic information intervention could augment impacts of information, saving costs by combining delivery of *models* together with contextual information. We plan on pursuing answers to these and other questions in future research.

6 Bibliography

Agwu, Agwu Ekwe, Murari Suvedi, Charity Chanza, Kristin Davis, Anges Oywaya-Nkurumwa, M Najjingo Mangheni, and PVK Sasidhar. 2023. "Agricultural extension and advisory services in nigeria, malawi, south africa, uganda, and kenya." *Partnerships for Innovative Research in Africa (PIRA) Research Report. East Lansing, Michigan, USA: Alliance for African Partnership, Michigan State University.*

Ahn, Woo-kyoung, and Charles W. Kalish. 2000. "The role of mechanism beliefs in causal reasoning." In *Explanation and cognition.* 199–225. Cambridge, MA, US:The MIT Press.

Ahn, Woo-kyoung, Charles W. Kalish, Douglas L. Medin, and Susan A. Gelman. 1995. "The role of covariation versus mechanism information in causal attribution." *Cognition*, 54(3): 299–352.

Alidaee, Hossein. 2023. "How Uncertainty About Heterogeneity Impacts Technology Adoption."

Amann, Arabel, Mathew Herrnegger, Jeninah Karungi, Allan John Komakech, Hope Mwanake, Lea Schneider, Christoph Schürz, Gabriel Stecher, Alice Turinawe, Matthias Zessner, et al. 2021. "Can local nutrient-circularity and erosion control increase yields of resource-constraint smallholder farmers? A case study in Kenya and Uganda." *Journal of Cleaner Production*, 318: 128510.

Anderson, Jock R., and Gershon Feder. 2007. "Chapter 44 Agricultural Extension." In *Agricultural Development: Farmers, Farm Production and Farm Markets.* Vol. 3 of *Handbook of Agricultural Economics*, , ed. R. Evenson and P. Pingali, 2343–2378. Elsevier.

Atuhaire, Aggrey, David Ocan, and Erik Jørs. 2016. "Knowledge, Attitudes, and Pratices of Tomato Producers and Vendors in Uganda." *Advances in Nutrition & Food Science*, 1(1): 1–7.

Banerjee, Abhijit V, and Esther Duflo. 2009. "The experimental approach to development economics." *Annu. Rev. Econ.*, 1(1): 151–178.

Banerjee, Abhijit V, Swati Bhattacharjee, Raghabendra Chattopadhyay, Esther Duflo, Alejandro J Ganimian, Kailash Rajah, and Elizabeth S Spelke. 2025. "Children's arithmetic skills do not transfer between applied and academic mathematics." *Nature*, 1–9.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on Treatment Effects after Selection among High-Dimensional Controls." *The Review of Economic Studies*, 81(2): 608–650.

Bold, Tessa, Kayuki C Kaizzi, Jakob Svensson, and David Yanagizawa-Drott. 2017. "Lemon technologies and adoption: measurement, theory and evidence from agricultural markets in Uganda." *The Quarterly Journal of Economics*, 132(3): 1055–1100.

Buyinza, J, J Sekatuba, H Agaba, R Kinuthia, and E Kiptot. 2015. "Analysis of Extension Systems in Uganda for Identification of Suitable Extension Approaches for Scaling-up "Trees for Food Security Project" in Eastern Uganda." *NaFORRI, Kampala, Uganda/ICRAF, Nairobi, Kenya*, 64.

Callander, Steven. 2011. "Searching and learning by trial and error." *American Economic Review*, 101(6): 2277–2308.

Callander, Steven, and Niko Matouschek. 2019. "The risk of failure: Trial and error learning and long-run performance." *American Economic Journal: Microeconomics*, 11(1): 44–78.

Chakraborty, Aranya. 2022. "Network-Based Targeting with Heterogeneous Agents for Improving Technology Adoption."

Chandrasekhar, Arun G., Esther Duflo, Michael Kremer, João F. Pugliese, Jonathan Robinson, and Frank Schilbach. 2022. "Blue Spoons: Sparking Communication About Appropriate Technology Use."

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Iván Fernández-Val. 2025. "Fisher–Schultz Lecture: Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, With an Application to Immunization in India." *Econometrica*, 93(4): 1121–1164. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA19303.

Craver, Carl, James Tabery, and Phyllis Illari. 2024. "Mechanisms in Science." In *The Stanford Encyclopedia of Philosophy*. . Fall 2024 ed., , ed. Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University.

CultivAid. 2021. "Mbale, Uganda Desk Review." Desk Review Report. Available at `https://www.cultivaid.org/wp-content/uploads/2021/09/Mbale-Uganda-Desk-Review-20210919.pdf`.

Davies, Benjamin, and Anirudh Sankar. 2025. "The value of conceptual knowledge."

Deaton, Angus. 2010. "Instruments, randomization, and learning about development." *Journal of economic literature*, 48(2): 424–455.

De Janvry, Alain, Karen Macours, and Elisabeth Sadoulet. 2017. *Learning for adopting: Technology adoption in developing country agriculture.* Ferdi.

Dewey, John. 1927. *The Public and Its Problems.* New York:Henry Holt Co. "An Essay in Political Inquiry".

Edaku, Charles. 2017. "The role of planning frameworks in supporting livelihoods activities in Uganda. A case of Mt Elgon, Bugisu region, Eastern Uganda." *Rural-Urban Dynamics in the East African Mountains*, 149.

Eliaz, Kfir, Simone Galperti, and Ran Spiegler. 2025. "False Narratives and Political Mobilization." *Journal of the European Economic Association.*

Esponda, Ignacio, Emanuel Vespa, and Sevgi Yuksel. 2024. "Mental models and learning: The case of base-rate neglect." *American Economic Review*, 114(3): 752–782.

Fisher, Matthew, and Frank C Keil. 2016. "The curse of expertise: When more knowledge leads to miscalibrated explanatory insight." *Cognitive science*, 40(5): 1251–1269.

Frazier, Brandy N, Susan A Gelman, and Henry M Wellman. 2009. "Preschoolers' search for explanatory information within adult–child conversation." *Child development*, 80(6): 1592–1611.

Gabaix, Xavier. 2014. "A sparsity-based model of bounded rationality." *The Quarterly Journal of Economics*, 129(4): 1661–1710.

Glennan, Stuart. 2017. *The new mechanical philosophy.* Oxford University Press.

Glover, Dominic, Jean-Phillipe Venot, and Harro Maat. 2017. "On the Movement of Agricultural Technologies: Packaging, unpacking, and situated reconfiguration." In *Agronomy for Development: The Politics of Knowledge in Agricultural Research. Pathways to Sustainability Series*, 14–30. New York, NY:Routledge.

Haeusler, Carole, and Jennifer Donovan. 2020. "Challenging the science curriculum paradigm: Teaching primary children atomic-molecular theory." *Research in Science Education*, 50(1): 23–52.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein. 2014. "Learning Through Noticing: Theory and Evidence from a Field Experiment*." *The Quarterly Journal of Economics*, 129(3): 1311–1353.

Hansson, Sven Ove. 2019. "Farmers' experiments and scientific methodology." *European Journal for Philosophy of Science*, 9(3): 32.

Havlin, J. L., J. D. Beaton, S. L. Tisdale, and W. L. Nelson. 1999. *Soil fertility and fertilizers: an introduction to nutrient management.* Upper Saddle River:Prentice-Hall Inc.

Hengl, Tomislav, Matthew A. E. Miller, Josip Križan, Keith D. Shepherd, Andrew Sila, Milan Kilibarda, Ognjen Antonijević, Luka Glušica, Achim Dobermann, Stephan M. Haefele, Steve P. McGrath, Gifty E. Acquah, Jamie Collinson, Leandro Parente, Mohammadreza Sheykhmousa, Kazuki Saito, Jean-Martial Johnson, Jordan Chamberlin, Francis B. T. Silatsa, Martin Yemefack, John Wendt, Robert A. MacMillan, Ichsani Wheeler, and Jonathan Crouch. 2021. "African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning." *Scientific Reports*, 11(1): 6130. Publisher: Nature Publishing Group.

Hoel, Jessica B., Hope Michelson, Ben Norton, and Victor Manyong. 2024. "Misattribution prevents learning." *American Journal of Agricultural Economics*, 1–24.

Islam, Mahnaz, and Sabrin Beg. 2021. "Rule-of-Thumb Instructions to Improve Fertilizer Management: Experimental Evidence from Bangladesh." *Economic Development and Cultural Change*, 70(1): 237–281. Publisher: The University of Chicago Press.

Jack, B Kelsey, Kathryn McDermott, and Anja Sautmann. 2022*a*. "Multiple price lists for willingness to pay elicitation." *Journal of Development Economics*, 159: 102977.

Jack, Kelsey, Kathryn McDermott, and Anja Sautmann. 2022*b*. "Multiple Price Lists for Willingness to Pay Elicitation."

Joo, Sehrang, Sami Ryan Yousif, and Frank Keil. 2021. "Understanding 'How': Two kinds of mechanistic explanations underlie known explanation preferences."

Jussupow, Ekaterina, Izak Benbasat, and Armin Heinzl. 2020. "Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion."

Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry. 2014. "Agricultural decisions after relaxing credit and risk constraints." *The Quarterly Journal of Economics*, 129(2): 597–652.

Keil, Frank C. 2006. "Explanation and understanding." *Annual Review of Psychology*, 57: 227–254.

Keil, Frank C. 2022. *Wonder: Childhood and the lifelong love of science.* MIT Press.

Kizza, Teddy, LosiraNasirumbi Winfred Nakyagaba, Nicolas Kigundu, Damalie B Magala, Monica Muyinda, Teopista Namirimu, and Dan Kirimani. 2017. "Irrigation technology utilization: A case of central Uganda."

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75(1): 83–119.

Laajaj, Rachid, and Karen Macours. 2024. "The Complexity of Multidimensional Learning in Agriculture."

Lattimore, Tor, and Csaba Szepesvári. 2020. *Bandit algorithms.* Cambridge University Press.

Levi-Strauss, Claude. 1966. *The savage mind.* University of Chicago Press.

Machamer, Peter, Lindley Darden, and Carl F Craver. 2000. "Thinking about mechanisms." *Philosophy of science*, 67(1): 1–25.

Maertens, Annemie, Hope Michelson, and Vesall Nourani. 2021. "How do farmers learn from extension services? Evidence from Malawi." *American Journal of Agricultural Economics*, 103(2): 569–595.

Magruder, Jeremy R. 2018. "An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries." *Annual Review of Resource Economics*, 10(Volume 10, 2018): 299–316. Publisher: Annual Reviews.

Matthias, Meg. 2025. "What Was Life Like Before We Knew About Germs?" *Encyclopædia Britannica – Story*, Retrieved August 16, 2025.

Miehe, Caroline, Leocardia Nabwire, Robert Sparrow, David Spielman, and Bjorn Van Campenhout. 2025. "Miracle seeds: Biased expectations, complementary input use, and the dynamics of smallholder technology adoption." *Economic Development and Cultural Change*. Publisher: The University of Chicago Press.

Ministry of Agriculture, Animal Industry and Fisheries, Republic of Uganda. 2016. "National Fertilizer Policy." Ministry of Agriculture, Animal Industry and Fisheries, Republic of Uganda. Policy document — Uganda National Fertiliser Policy (August 2016).

Ministry of Agriculture, Animal Industry and Fisheries Uganda. 2016. "Guidelines and Standards for Agricultural Extension and Advisory Services in Uganda." Ministry of

Agriculture, Animal Industry and Fisheries. FAOLEX document UGA/223402, available at https://faolex.fao.org/docs/pdf/uga223402.pdf.

Mokyr, Joel. 2016. "A culture of growth: The origins of the modern economy."

Munshi, Kaivan. 2004. "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution." *Journal of Development Economics*, 73(1): 185–213.

Nourani, Vesall, Nava Ashraf, and Abhijit Banerjee. 2025. "Learning to Teach by Learning to Learn."

Nyaga, Justine M, Cecilia M Onyango, Johanna Wetterlind, and Mats Söderström. 2021. "Precision agriculture research in sub-Saharan Africa countries: A systematic map." *Precision Agriculture*, 22(4): 1217–1236.

Perez, Jasmin, and Lisa Feigenson. 2022. "Violations of expectation trigger infants to search for explanations." *Cognition*, 218: 104942.

Poulton, Colin, Jonathan Kydd, and Andrew Dorward. 2006. "Increasing fertilizer use in Africa: What have we learned." *Agriculture and Rural Development Discussion Paper*, 25: 43.

Rozenblit, Leonid, and Frank Keil. 2002. "The misunderstood limits of folk science: An illusion of explanatory depth." *Cognitive science*, 26(5): 521–562.

Schwartzstein, Joshua. 2014. "Selective attention and learning." *Journal of the European Economic Association*, 12(6): 1423–1452.

Schwartzstein, Joshua, and Adi Sunderam. 2021. "Using Models to Persuade." *American Economic Review*, 111(1): 276–323.

Sen, Amartya. 1999. *Development as Freedom.* New York:Oxford University Press.

Spiegler, Ran. 2020. "Behavioral Implications of Causal Misperceptions." *Annual Review of Economics*, 12(1): 81–106. _eprint: https://doi.org/10.1146/annurev-economics-072219-111921.

Sserunjogi, Brian, Florence Nakazi, Sheila Nakkazi, and Rehema Kahunde. 2023. "Public Funding for Agricultural Extension Services in Uganda: Potential Sources for Budget Reallocation for the Financial Year 2023/24." Economic Policy Research Centre Policy Note 16. Available at https://eprcug.org/publication/public-funding-for-agricultural-extension-services-in-uganda-potential-sources-for-budget-real

Stahl, Aimee E, and Lisa Feigenson. 2015. "Observing the unexpected enhances infants' learning and exploration." *Science*, 348(6230): 91–94.

Suri, Tavneet, and Christopher Udry. 2022. "Agricultural Technology in Africa." *Journal of Economic Perspectives*, 36(1): 33–56.

Tjernström, Emilia. 2017. "Learning from others in heterogeneous environments." *Unpublished manuscript*.

Tjernström, Emilia, Travis J. Lybbert, Rachel Frattarola Hernández, and Juan Sebastian Correa. 2021. "Learning by (virtually) doing: Experimentation and belief updating in smallholder agriculture." *Journal of Economic Behavior & Organization*, 189: 28–50.

Vos, Rob, Joseph Glauber, Charlotte Hebebrand, and Brendan Rice. 2025. "Global shocks to fertilizer markets: Impacts on prices, demand and farm profitability." *Food Policy*, 133: 102790.

Vygotsky, L. S. 1978. *Mind in Society: Development of Higher Psychological Processes.* . 14th ed., Harvard University Press.

Wacal, Cosmas, Daniel Basalirwa, John Byalebeka, Mitsuri Tsubo, and Eiji Nishihara. 2023. "Low cost maize stover biochar as an alternative to inorganic fertilizer for improvement of soil chemical properties, growth and yield of tomatoes on degraded soil of Northern Uganda." *BMC Plant Biology*, 23(1): 473.

World Bank. 2025. "Employment in agriculture (% of total employment) (modeled ILO estimate) – Uganda." *https://data.worldbank.org/indicator/SL.AGR.EMPL.ZS?locations=UG*, Accessed: August 14, 2025.

Zajonc, Robert B. 1968. "Attitudinal effects of mere exposure." *Journal of personality and social psychology*, 9(2p2): 1.

A Appendix: Secondary Outcomes, and Additional Tables & Figures

A.1 Secondary Outcomes

Diagnosis and Treatment Respondents complete two exercises pertaining to diagnosing deficiencies and recommending improvements. In the first exercise, we present farmers with a vignette of a tomato

plant that has grown vigorously with a strong stem but produces fruits of dull and uneven color. Respondents are asked to choose, from four fertilizer recipes, which the farmer most likely applied. Three options provide substantial potassium, while one provides minimal potassium.

In our framework, this task asks whether respondents can invert the mapping from outcomes to inputs. Given an observed result y, can they infer a likely input bundle by first mapping outcomes to nutrients via g^{-1}, and then nutrients to inputs via z^{-1}, i.e., $x \in z^{-1}(g^{-1}(y))$? In particular, for a poor result y, the task tests whether respondents behave as if they identify a culprit nutrient vector $z = g^{-1}(y)$.

This exercise captures diagnostic reasoning. Farmers frequently observe plant symptoms and must infer which nutrients are missing. Without mechanistic understanding, they may rely on superficial cues or fail to link symptoms to underlying nutrient dimensions, leading to ineffective adjustments.

In a second exercise, respondents are presented with a farmer who applies a fertilizer recipe consisting only of Urea. They are asked to recommend an addition to improve the recipe. Three options are primarily nitrogen-based, while one option (NPK 17-17-17) provides substantial amounts of all three macronutrients.

In terms of our framework, we test whether farmers attend to the principle of limiting factors (Havlin et al., 1999), corresponding to a Leontief-like structure in $g(z)$. Farmers reasoning mechanistically should recognize that primarily adding more nitrogen may have low returns if nitrogen is not the limiting (bottleneck) nutrient, and should instead consider inputs that address other nutrient constraints.

This exercise captures treatment reasoning. Farmers often seek to improve an existing fertilizer plan, and mechanistic understanding provides a compass for profitable experimentation in a combinatorially vast space of possible input combinations. In this context, it guides which adjustments address binding constraints rather than redundantly increasing already-supplied nutrients.

Table A1 presents the results from these two exercises. In Column (2), the outcome variable is a dummy for whether the respondent correctly selects the recipe that would explain the nutrient deficiency in the tomato plant. The effect size is small and insignificant. The mechanisms treatment did not help farmers attribute the observed symptoms to the likely nutrient deficiency. We now turn to Column (3), where the outcome variable indicates whether the respondent chose the recipe most likely to address a limiting factor in the vignette for the farmer who only applies Urea. Here the effect of the treatment is positive, large, and significant. Farmers in the treatment group are 13.8 percentage points more likely to choose the correct adjustment, an increase of 32% from the control mean ($p < 0.01$). This suggests that mechanistic explanations improve performance on a realistic choice problem: adjusting levels of an input from a suboptimal baseline.

Table A1: Nutrient Deficiency Problems

	(1) Index	(2) Diagnosing	(3) Limiting Factor
Treat	0.170	0.030	0.138
	(0.051)	(0.032)	(0.035)
	[0.001]	[0.353]	[0.000]
Observations	797	797	797
Control Mean	1.87e-08	0.287	0.431
Control SD	0.736	0.453	0.496
Mean (unstandardized)	0.359		

Recipe Translation Respondents are shown results from the fertilizer demonstration conducted on grey soil and asked to predict yields in three related contexts: grey soil without manure, black soil with manure, and black soil without manure. Predictions are elicited sequentially, and after each response, respondents observe the agronomist benchmark before proceeding to the next prediction.

In our framework, respondents observe a return in one context, $g(z', k) - g(z, k) = \delta$, and are asked to extrapolate to a new context k'. We test whether they use structural relationships across contexts, such as monotone or increasing differences, to infer $g(z', k') - g(z, k')$ from observed signals.

This exercise captures the ability to translate information across heterogeneous environments. Farmers often observe returns in one soil or management context and must infer applicability elsewhere. Mechanistic understanding of nutrient retention and soil properties may allow farmers to extrapolate returns across contexts rather than treating each setting as unrelated.

The results are shown in Table A2. First in Column (1) we look at impacts on the difference in predicted yields for the demo recipe with manure on the black soil and the grey soil. The impact of the treatment is null here, potentially because the control mean of this predicted difference is already quite high: the average control group respondent expects yields 67% higher on the black soil. Column (2) shows results from the extrapolation problem described above. After a farmer learns the returns to manure on grey soil, how do they update their beliefs about the returns to manure on black soil? In the control group, the average predicted difference is negative, the opposite sign of the relationship predicted by agronomic principles. The treatment increases predicted returns to manure on grey soils relative to black soils by 7.7 percentage points, an effect that is statistically significant at the 10% level ($p = 0.071$). Mechanistic explanations allowed farmers to draw inferences from returns in one context (yield gains from manure on grey soil) to another, correcting baseline beliefs which held the opposite sign. This evidence suggests that an understanding of the mechanisms behind why returns differ by context may help farmers to learn in contexts with heterogeneous returns, a classic problem in social learning that has been documented since

Munshi (2004).

Table A2: Extrapolation Across Soil Types

	(1) Black vs. Grey	(2) Manure Returns (Grey vs. Black)
Treat	-0.008	0.077
	(0.075)	(0.043)
	[0.911]	[0.071]
Observations	797	797
Control Mean	0.673	-0.134
Control SD	1.148	0.733

Meta-mechanisms Respondents are presented with four pairs of opposing statements about fertilizer use and asked to indicate which they agree with. The statements address context dependence, timing versus quantity, differences between synthetic fertilizers and manure, and soil fertility management. In each pair, one statement is an implication of mechanistic understanding, while the opposing statement is mechanistically unsupported. Importantly, these implications are not explicitly included in the mechanisms training, but are simple deductive consequences of it. The paired statements are:

1. **Context**: (a) *Which fertilizer is better depends on the specific scenario in which it will be used*; (b) *Between NPK 17-17-17 and DAP, one fertilizer is better than the other overall*

2. **Timing**: (a) *When using fertilizers, the best way to increase yields is to apply the right fertilizers at the right times, regardless of how much you apply*; (b) *When using fertilizers, the best way to increase yields is to increase the amount of fertilizer you apply, regardless of when you apply it*

3. **Synthetic**: (a) *Chemical fertilizers and manure do some things differently things to the soil, no matter how much or which fertilizers you apply*; (b) *Chemical fertilizers and manure both do exactly the same thing to the soil: they only differ in the amounts you need to apply to achieve the desired result*

4. **Soil Health**: (a) *There are fertilizer practices that can improve a soil's fertility*; (b) *There is little that one can do about declining fertility of the soil*

In each pair, (a) is an implication of mechanistic understanding while (b) is mechanistically unsupported. These implications can be thought of as "meta-mechanisms." Psychology suggests that even when detailed mechanisms are forgotten, key implications may persist as a conceptual scaffolding (Keil,

2006). In the case of fertilizers, such scaffolding can meaningfully guide choice. For example, recognizing that manure and synthetic fertilizers play different roles may make it more likely to deploy them as complements rather than substitutes.

Table A3 presents the results from these questions. In each column, the outcome variable is an indicator variable for whether the respondent somewhat or strongly agreed with statement (a) over statement (b). To reduce the impact of order effects, the order of presentation for statement (a) and (b) is randomized for each question. Column (1) shows results for a standardized index of the four dummies. Overall, the treatment boosts agreement with the correct rule of thumb. The standardized index increases by 0.06 points for the treatment group, though it is just above the threshold of significance ($p = 0.105$). These effects are concentrated on the *Timing* statement, and to a lesser extent *Soil Health*. The point estimates for the other two statements, *Context* and *Synthetic*, are positive but small and insignificant. Overall, the results suggest that mechanistic explanations weakly impart mechanistic metadata that can guide choice, even when these rules are not explicitly taught or emphasized.

Table A3: Meta-mechanisms

	(1) Index	(2) Context	(3) Timing	(4) Synthetic	(5) Soil Health
Treat	0.057	0.022	0.049	-0.001	0.049
	(0.035)	(0.034)	(0.027)	(0.035)	(0.033)
	[0.105]	[0.531]	[0.070]	[0.980]	[0.140]
Observations	797	797	797	797	797
Control Mean	-6.35e-09	0.490	0.782	0.480	0.530
Control SD	0.514	0.501	0.414	0.500	0.500

Psychological Attitudes An understanding of the mechanistic explanations behind a phenomenon may be psychologically empowering, apart from its instrumental utility. A person who understands how a technology works may feel less intimidated in operating or experimenting with it, may have a greater sense of self-efficacy in relation to the technology, and feel less fearful of negative or unintended consequences. These psychological attitudes are subtle but may be important drivers of behavioral change. To test whether psychological attitudes toward a technology shift after exposure to mechanistic explanations, we ask farmers a series of questions about their attitudes toward using fertilizers. These cover attitudes toward experimentation, expectations of yield increases, and perceptions of the risk of crop burning and toxicity (two commonly cited fears associated with fertilizer use).

The results from these questions are displayed in Table A4. In each column, the outcome variable is scaled so that a positive sign indicates positive affect (greater locus of control, optimism, less fear of

negative consequences). Outcomes in Columns (2), (4), and (5) are dummy variables for whether the respondent somewhat or strongly agrees with the positive attitude. The outcome in Column (3) is the percentage point increase in yield the farmer believes he can achieve in the next season. Effects on all of these questions are negative but small and statistically insignificant. However, uniformly high control means—such as the 74% expected yield increase in Column (3)—suggest substantial experimenter demand effects. We therefore find no evidence of changes in psychological attitudes, though the elicitation may itself be compromised by demand effects.

Table A4: Psychological Attitudes

	(1) Index	(2) Can Exper.	(3) Boost Yield	(4) Not Burn	(5) Not Toxic
Treat	-0.028	-0.019	-1.514	-0.009	-0.004
	(0.036)	(0.034)	(1.395)	(0.026)	(0.026)
	[0.427]	[0.572]	[0.278]	[0.725]	[0.875]
Observations	797	797	797	797	797
Control Mean	-4.79e-09	0.617	74.22	0.843	0.822
Control SD	0.513	0.487	22.75	0.365	0.383

A.2 Additional Tables

Table A5: Best linear predictor of CATE using (Chernozhukov et al., 2025)

	ATE	HET	ATE	HET
	Random Forest		Elastic Net	
Profits	20.65	0.596	22.57	0.265
	(4.267,37.29)	(0.383,0.812)	(2.488,42.20)	(0.095,0.434)
	[0.077]	[0.000]	[0.130]	[0.022]

Table A6: Goodness-of-fit criteria of (Chernozhukov et al., 2025)

	Elastic Net	Random Forest
Profits	1789.00	6029.00
	29.96	102.70

Table A7: CLAN of nonzero difference in baseline var between top/ bottom GATE quintile, i.e. $E[\text{var}, G_5] - E[\text{var}, G_1] \neq 0$, from random forest ML proxy estimation using (Chernozhukov et al., 2025)

	Top Quintile	Bottom Quintile	Difference
$\boldsymbol{E[\text{var}, G_5] - E[\text{var}, G_1] > 0}$			
Completed High School	0.311	0.181	0.128
	(0.266,0.357)	(0.136,0.226)	(0.065,0.192)
	-	-	[0.002]
Literate	0.654	0.488	0.165
	(0.603,0.706)	(0.436,0.539)	(0.092,0.238)
	-	-	[0.000]
Knowledge Index	0.039	-0.015	0.056
	(0.004,0.075)	(-0.051,0.021)	(0.005,0.107)
	-	-	[0.142]
Use Irrigation Pump	0.105	0.029	0.074
	(0.079,0.131)	(0.003,0.055)	(0.038,0.111)
	-	-	[0.001]
Experimented with New Fertilizer	0.189	0.116	0.074
	(0.151,0.227)	(0.078,0.155)	(0.021,0.127)
	-	-	[0.044]
Use MOP	0.025	0.000	0.021
	(0.013,0.036)	(-0.011,0.013)	(0.006,0.038)
	-	-	[0.049]
Use Winner	0.095	0.025	0.070
	(0.070,0.119)	(0.000,0.049)	(0.035,0.106)
	-	-	[0.002]
Use Power	0.012	0.000	0.012
	(0.004,0.021)	(-0.008,0.008)	(0.001,0.024)
	-	-	[0.167]
$\boldsymbol{E[\text{var}, G_5] - E[\text{var}, G_1] < 0}$			
Years Grown Tomatoes	9.043	11.27	-2.119
	(8.094,10.02)	(10.32,12.27)	(-3.447,-0.769)
	-	-	[0.020]
Pct Farm Income from Tomatoes	54.46	59.99	-5.748
	(51.86,57.01)	(57.45,62.58)	(-9.391,-2.106)
	-	-	[0.018]
No. Tomato Plots	1.241	1.464	-0.222
	(1.161,1.321)	(1.385,1.543)	(-0.331,-0.117)
	-	-	[0.002]
Tomato Yield	4.893	5.785	-0.899
	(4.309,5.497)	(5.173,6.384)	(-1.735,-0.096)
	-	-	[0.132]
Brown Soil	0.143	0.221	-0.081
	(0.103,0.184)	(0.180,0.262)	(-0.137,-0.023)
	-	-	[0.041]
Prefer 15-0 Lottery Over 10	0.306	0.386	-0.082
	(0.256,0.356)	(0.336,0.436)	(-0.152,-0.011)
	-	-	[0.115]
Prefer 20-0 Lottery Over 10	0.316	0.421	-0.102
	(0.266,0.367)	(0.370,0.472)	(-0.174,-0.031)
	-	-	[0.038]
Prefer 25-0 Lottery Over 10	0.336	0.426	-0.095
	(0.285,0.387)	(0.376,0.477)	(-0.167,-0.023)
	-	-	[0.058]
Prefer 30-0 Lottery Over 10	0.349	0.453	-0.104
	(0.297,0.401)	(0.401,0.505)	(-0.176,-0.031)
	-	-	[0.037]
Prefer 35-0 Lotter Over 10	0.336	0.439	-0.100
	(0.285,0.387)	(0.388,0.489)	(-0.173,-0.027)
	-	-	[0.048]

Table A8: Endline: Marginal effect of Reinforcement Training on Input Choices

	(1) % K fruiting	(2) Max N concentration
Treat	3.797	-3.072
	(2.752)	(1.552)
	[0.168]	[0.048]
Treat*Reinforce	1.412	-0.259
	(4.098)	(2.164)
	[0.730]	[0.905]
Observations	970	970
Control Mean	25.55	71.04
Control SD	35.73	20.48

Table A9: Endline: Marginal Effect of Reinforcement Training on Yields

	(1) Yields per Acre	(2) Yields per Acre
Treat	0.910	0.977
	(0.520)	(0.512)
	[0.080]	[0.056]
Treat*Reinforce	-0.107	-0.161
	(0.821)	(0.820)
	[0.896]	[0.845]
Control Shocks	No	Yes
Observations	1000	1000
Control Mean	6.370	6.370
Control SD	6.209	6.209

A.3 Additional Figures

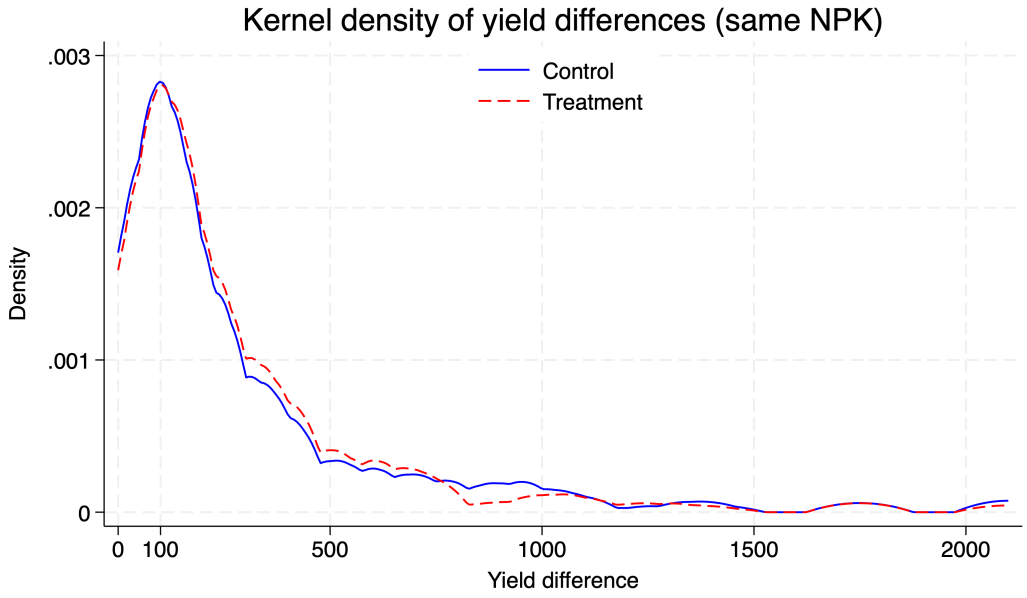


Figure 15: Kernel density of reported differences between macronutrient equivalent recipes

B Appendix: Treatment and Control Trainings

TREATMENT TRAINING

The extension officer was trained by our soil-scientist coauthor on the curricula below for treatment and control trainings.

What is Soil Made of?
Soil is an important medium for plant growth and the way we nourish it largely determines how we benefit from it.

- Soil is made up of five basic components: **minerals**, **organic matter**, water and air, living organisms.
- Minerals are essentially fine particles of stone
- Organic matter is essentially dead living things (plants, animals, microbes)

4.1 Soil texture

- The mineral portion of soil is divided into three classes/types: **sand, silt, and clay**
    - Sand, Silt, and Clay are mineral ('stone') particles of different **sizes**
    - Sand is very large, silt is medium to large, and clay is very small
        - *Show a ball which is 25x larger than a smaller ball - this is the size difference of clay to the smallest sand particle*
    - They are all mixed up together in ALL soils, but the AMOUNT of clay, sand, and silt differs
- The relative proportion of sand, silt, or clay in a soil is known as soil texture
    - Introduce buckets of clay rich, sand rich, and loamy soils – feel the difference
    - Smaller particles stick to things better – like water, organic matter, and nutrients
        - Larger particles don't stick to things very well
        - This 'stickiness' is the main thing to understand how manage different soil textures
- Soil texture is VERY IMPORTANT for crop production and many aspects of soil management
    - More clays = More small sized particles = More stickiness
    - More sand = More large sized particles = Less stickiness
    - Ability to store soil nutrients:
        - Sandy soils have
            - Low plant nutrients naturally in them
            - Low capacity to retain applied nutrients
            - Low ability to hold water (which is important for nutrients to be taken up by the plant)
        - Clay soils have
            - High plant nutrients naturally in them
            - High capacity to retain applied nutrients
            - High ability to hold water (which is important for nutrients to be taken up by the plant)
        - Too much sand: Poor fertility, low water holding, but easy to dig
        - Too much clay: High nutrient storage, high water holding, but very hard and sticky
        - Good mix of sand and clay is what we call LOAMY soils.

1.2 Soil Fertility Management

4.2.1 Plant nutrients

- A plant requires 16 '**nutrients'**
- A nutrient is a specific '**element'**, which are the building blocks of nature
- In addition to the many others, the plant needs 3 key nutrients in high quantity which it can only get from the soil
    - **Nitrogen**, **Phosphorus**, **Potassium**
- Soil nutrients required by the plant are present in form of molecules.
    - **Molecules** are just groups of elements that are stuck together
    - The plant absorbs almost all these soil nutrients through the roots
- Nutrients come mainly from 3 sources
    - From the minerals / (tiny rocks) in the soil
    - From the organic matter (dead things) in the soil
    - From fertilizer (synthetic or organic) applied by people
- Some nutrients may be in the soil but not be available to the plant
    - This is because they can be 'stuck' to the soil particles
- **Synthetic** and **Organic** sources of fertilizer
    - Organic fertilizers are sources of nutrients that are made of dead / living things
        - Examples are manure (dead grass and microbes, compost (decomposed plant materials)
        - Manure has some nutrients but not in concentrated form, and only becomes available to the plant slowly
        - Manure has many positive effects on the soil, making it better for crop production in ways that are not related to nutrients.
            - This includes improving structure, nutrient holding ability, water holding capacity, buffering acidity
    - Synthetic fertilizers supply the plant nutrients in forms that are immediately available to plants
        - They are the most concentrated form of nutrients and are applied in much smaller quantity than organic fertilizer
        - They don't improve other features of the soil, they just supply plants with their immediate needs for nutrients
        - They made from common parts of our world, nitrogen is made from the air, phosphorus and potassium are mined from rocks and refined into more concentrated forms
    - No matter whether fertilizer is applied as organic or synthetic forms, they both break down into the same molecules which are used by the plant

1.2.2 Major plant nutrients and their availability in the soil

Nitrogen

What it is:

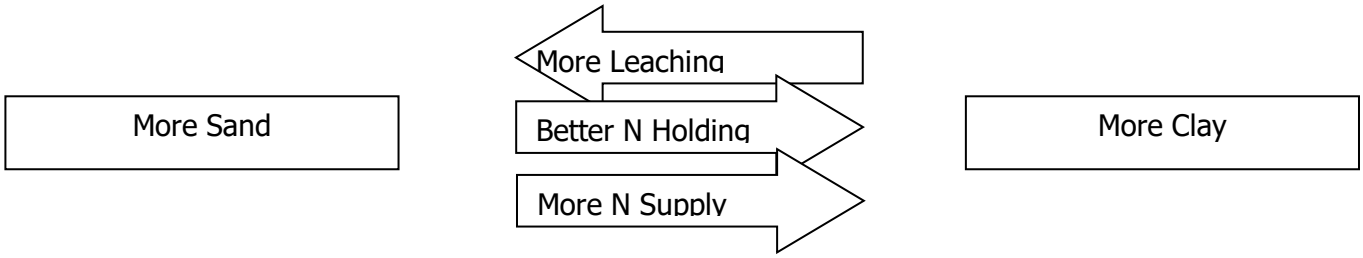
- It is a key element in improving plant growth vigor and greenness.
- It is the soil nutrient needed in largest quantity by the plant

How Nitrogen Behaves in the Soil

- Most of the nitrogen in soil is in the organic matter part of the soil, there is not so much in the mineral part of the soil
- When there is water in the soil, nitrogen is released from the organic matter in a form that plants can take up
- Once released, soils have limited capacity to keep the nitrogen; it must be taken up by plants or

it is lost
- Water in the soil is necessary for plants to take up nitrogen
- When there is too much water passing through the soils, nitrogen can easily be washed deep into the soil where plants cant get it (eg when there is heavy rain) – This is called **leaching**
- When there is too much water stagnating in the soil, microorganisms in the soil convert the nitrogen to gas which the plants can't use and is lost to the air

How visible soil features affect Nitrogen Availability in Soil
- Soil Organic matter / color:
    - Darker soils tend to have organic matter and therefore more N naturally released in the soil
    - Organic matter creates more places for N to bind to soil particles and reduces leaching
- Soil texture and mineralogy:
    - Sandier soils (gritty or coarse soils) have lower capacity to supply nitrogen because they generally hold less organic matter
    - Sandier soils have fewer places for N to bind to soil particles
    - Sandier soils are unlikely to be water-logged for a long time, so have less risk of losing N to the air
    - Sandier soils tend allow water to pass through more quickly which leads to greater risk of losing N to leaching
- Flooding/Deposition
    - Places with flooding / deposition once every few years or more receive a lot of organic matter additions
    - In these cases, the N supply can be higher than other similar soils that don't receive flooding
    - This can be true even on sandy soils or soils that don't look very dark



Phosphorus

<u>What it is:</u>
- It is a key element in helping the plant form substances that help it get energy from the sun, and nutrients from the soil.
- It is a soil nutrient needed in large quantity by the plant
- It is especially important when the plant is still small, as it can be recycled many times within the plant

<u>How Phosphorus Behaves in the Soil</u>
- Phosphorus in soil is stored in the soil clay and organic matter
- Organic matter in the soil also helps release phosphorus from the clay in the soil
- As the humus decomposes, it releases phosphorus in a form that plants can take up
- Phosphorus is taken up by plant roots when dissolved in water. Water in the root zone is therefore necessary for plants to take up phosphorus.
- Phosphorus is released very slowly into the soil water by clay and organic matter,
- Because it is released very slowly, phosphorus is not prone to leaching, unlike nitrogen.
- Phosphorus does not easily form into gas that can be lost from the soil so it is not affected by

water-logging conditions, unlike nitrogen.

- The clay and organic matter particles that hold phosphorus are easily carried away during soil erosion.
- Therefore, where the land is prone to soil erosion (steep slopes, bare soil) large amounts of phosphorus can be lost if not controlled through practices such as mulching.

How visible soil features affect Phosphorus Availability in Soil

- Soil Organic matter / color:
  - Red colored soils tend to have a high tendency to lock up phosphorus in forms that cannot be accessed by plants,
  - This causes the phosphorus in the soil to be available for the plant to use
  - Soils that are rich in organic matter need a smaller dose of phosphorus
    - Soils with more organic matter tend to be darker than similar nearby soils
- Soil texture and mineralogy:
  - Sand particles have low capacity to supply phosphorus, but clay particles have more ability to supply phsophrus
- Drainage
  - There is no need to alleviate water-logging to enhance phosphorus availability to plants.
- Flooding/Deposition
  - Places with flooding / deposition once every few years or more receive a lot of organic matter additions
  - In these cases, the N supply can be higher than other similar soils that don't receive flooding
  - This can be true even on sandy soils or soils that don't look very dark

Potassium

What It Is:

- It facilitates the manufacture of food in the leaves,
- regulates water loss from the plant,
- Supports the filling up of fruits, grains and tubers

How Potassium Behaves In The Soil

- Potassium in soil is stored in the clay and organic matter, which are like magnets attracting and holding onto the potassium
- Water in the root zone is key for plants to take up potassium.
- Potassium is released into soil water from the clay and humus particles, from which plants take it up and utilize it.
- Potassium is not as prone to leaching loss from the soil as nitrogen but can still be lost if the soil has limited amounts of clay and organic matter in it to hold the potassium that is in the soil water.
- It does not easily turn into gas that can be lost from the soil, so is not affected by water-logging, unlike nitrogen.
- The clay and organic matter particles that bear potassium are easily carried away during soil erosion.
- Therefore, where the land is prone to soil erosion (steep slopes, bare soil) large amounts of potassium can be lost if not controlled through practices such as mulching.

How Visible Soil Features Affect Potassium Availability In Soil

- Soil Organic matter / color:
  - Soils that are rich in humus, recognized from their dark brown color need a smaller dose of potassium from mineral fertilizers than that for light-colored soils
- Soil texture and mineralogy
  - Sandy soils tend to have low capacity to supply potassium.

- Soils rich in cracking clays need a smaller dose of potassium from mineral fertilizers than that for non cracking
- Drainage
    - There is no need to alleviate water-logging to enhance potassium availability to plants.

Why there is a need to manage fertilizer applications well

- Some fertilizer companies say a certain fertilizer is 'good for tomatoes' or 'good for maize'
    - This is like saying Peter likes posho
    - Even if he likes posho, whether Peter wants to eat posho depends on his situation at the time
    - If Peter has a lot of posho on his plate, but no sauce, he won't eat more posho even if you give him more
    - If you want him to eat more food, you also need to give him some sauce
    - This is also how plants take up nutrients, some might be enough in the soil while others are lacking, and the plant needs them all to continue growing well
    - So you need to consider what nutrients are in your soil, and how you add different nutrients together rather than look at which crops 'like' which fertilizer
- If you don't carefully select the fertilizer, you apply too much of one nutrient and too little of another
- This means you spend money on fertilizer you don't need, and the plant ends up lacking what it needs
- If you apply fertilizer at the wrong time, much of it may not get used by the plant

Practices That You Can Use To Manage Fertilzer

There are three ways you can manage fertilizer to get the most from it.
First, you can **select the right type** of fertilizer.
Second you can **select the right amount** of fertilizer
Third, you can **select the right timing** to apply the fertilizer.

- Form of N applied – Selecting the Type of Fertilizer
    - Fertilizers are labeled with the amount of N, P, and K in them. It is important to know which fertilizers are high in N, P, and K.
        - We can see that the high N fertilizers are Urea, CAN, and NPK 25:5:5
        - High P fertilizers are DAP,TSP, and NPK 17:17:17
        - High K fertilizer is MOP and NPK 17:17:17
        - Other fertilizers are sold but it is important to first check how much nutrient they have as some fertilizers are expensive but may have lower amounts of nutrients than these fertilizers

Fertilizer	Percent N	Percent P (P_2O_5)	Percent K (K_2o)
Urea	46	0	0
DAP	18	46	0
MOP	0	60	0
CAN	27	0	0
TSP	0	46	0
NPK (25:5:5)	25	5	5
NPK (17:17:17)	17	17	17

    - You can supply relatively doses of nitrogen from nitrogen-rich fertilizers, including urea and CAN, and NPK (25:5:5)
        - Other fertilizers have N in smaller doses, so you'd have to apply more kilos if the goal is to apply N

    - You can supply relatively doses of nitrogen from phosphrus-rich fertilizers, including DAP, TSP, and NPK (17:17:17)
        - Other fertilizers have P in smaller doses (like NPK 25:5:5), so you'd have to apply more kilos if the goal is to apply P
    - You can supply relatively doses of nitrogen from potassium-rich fertilizers, including MOP, and NPK (17:17:17)
        - Other fertilizers have K in smaller doses (like NPK 25:5:5), so you'd have to apply more kilos if the goal is to apply K
        -

- <u>The Amount of Fertilizer to Apply</u>
    - You can adjust how much of each fertilizer to apply, depending on the features of the soil you have
    - You can apply animal manures to supply some nitrogen and phosphorus, and improves the soil's capacity to store water and reduce leaching loss of nitrogen
    - You can consider the color as an indicator of organic matter which may indicate less or more need for N and P.
        - Darker soils tend to have more organic matter, compared to similar nearby soils which are lighter but otherwise share similar characteristics like texture

- <u>Timing of fertilizer applied</u>
    - Finally, you can adjust when to apply the fertilizer according to the features of the soil you have
    - For some soils it is ok to apply nitrogen in fewer applications but larger quantities because these soils can hold nitrogen better and are less likely to lose nitrogen
    - In other soils which lose nitrogen more easily it is more important to break up the dose of the nitrogen into 2 to 3 portions to minimize loss of the nitrogen.
    - For Phosphorus, the important thing to remember about timing is that the plant needs it early in its growth
        - Timing of phosphorus doesn't depend much on the soil type but mostly on the plant stage
    - For potassium, you may consider applying in more smaller doses than fewer bigger doses depending on the tendency of the soil supply and to lose K.

Fertilizer Applied in the Demo

- At Transplanting we applied
    - Manure at 1 tampecos if it is dry or 2 tampecos if it is wet per hole
    - NPK 17:17:17 1/3 glass bottle cap per hole
    - DAP at 1 ¼ glass bottle caps per hole
    - MOP at 1/3 glass bottle cap per hole
- We mixed all the manure and fertilizers with the soil in the hole until well mixed before planting the seedling

- At 3 weeks after planting we applied the first top dressing
    - DAP (g/plant): ¾ glass bottle caps per hole
    - Urea (g/plant): 1 glass bottle caps per hole
- Top dressing was done by digging a small/shallow hole between 10-15cm from the plant stem, placing the fertilizers into the hole, and covering with soil.

- At 8-9 weeks after transplanting, at full flower bloom, we applied the second top dressing with the following fertilziers
    - MOP (g/plant): ¾ glass bottle caps per hole
    - Urea (g/plant): ¼ glass bottle caps per hole
- Top dressing was done by digging a small/shallow hole between 10-15cm from the plant stem,

placing the fertilizers into the hole, and covering with soil.
- We decided this type, amount, and timing of fertilizer because we saw this was a sandy, light colored soil. Because of that we took the ability of the soil to naturally supply nutrients into account as well as the tendency for it to lose nutrients.

Review and Discussion
- What is soil made of?
- What is the meaning of texture ?
    - What is sand?
    - What is clay?
- What is organic matter?
- What are the three important nutrients we've talked about today?
    - Where do they come from in the soil?
    - How do they act differently in the soil?
    - What fertilizers are high in N?
    - What fertilizers are high in P?
    - What fertilizers are high in K?
- What practices can we use to apply N, P, and K fertilizers so they are most effective?

CONTROL TRAINING

Before starting these topics, besides the fertilizer recipe, the extension officer goes through the fertilizer bags without sharing the mechanisms. On the topics below, the extension officer goes into further depth on subsets of these per audience interest. Staking and pestcide use were popular topics

Other Aspects of Crop Management

Digging Planting Holes and Marking Sub-Plots

-Tie ropes to two sticks and measure out 10m of rope length between the sticks
-Start at edge 1 of the plot
-Dig planting holes every 60cm
    -Use a measured stick to measure spacing between holes
-Measure 90cm distance from first line to second line at both ends of the rope
-Continue digging planting holes in second line
-Complete digging for 20 rows
-Leave a 1 meter space between row 20 and the next row
-Continue planting with the same spacing for another 20 rows

Weed Control

-Throughout the season, weeds should be controlled at less than 4 leaf stage. This usually translates into weeding every 3-4 weeks.
-Mulching should be done if materials are available, but must be done to both plot A and B if done at all

Staking

-Staking should be done between 4-8 weeks after transplanting.
-Strong stakes should be collected at the plot, and main stems tied to one stake each
-As growth proceeds, additional branches / secondary stems should be tied progressively higher up the stake
-Each plant should be tied at least 3 times to the stake by full fruiting

Crop Protection

-Scouting for pest damage should be done at least twice per week
    -The host farmer may scout and provide reports to the demo plot manager, the plot manager should scout personally at least once per week
-Preventative Fungicide should be applied once every 5-7 days, along with a surfactant
    -Mancozeb is preferred, but clorothoranil is also acceptable
    -After flowering and first fruiting, a weekly rotation between Metalaxyl, Cymoxanil, or other systemic fungicides effective on early & late blight should be applied instead of Mancozeb on a 5-7 day interval
-At transplanting, apply a broad spectrum insecticide on the soil and plant around the stem to protect against cutworms
    -Lambda-cyhalothrine, cypermetherine, deltamethrine are acceptable insectides
    -Apply with a surfactant
-After transplanting scouting for the following pests can require the following insecticides:
    -Whiteflies: If present in excessive numbers, a systemic insecticide such as a acetamiprid or Azadirachtin if not many present, control is not needed
    -Thrips: If present in excessive numbers, a systemic insecticide such as a Abamectin, or dimethoate . if not many present, control is not needed
    -Bollworms: If present in excessive numbers, a broad spectrum insecticide should be applied such as Lambda-cyhalothrine, cypermetherine, deltamethrine
    -Tutta Absoluta: If any is seen in the plot, preventative application of either Emamectin benzoate or Flubendiamide
    -Red Spider Mites : If any is seen in the plot, a systemic insecticide such as a Abamectin, if not many present, control is not needed
-Apply all the above insecticides with a surfectant

C Appendix: Agronomic Model

We worked with an agronomist (co-author Godfrey Taulya) to develop a model for predicting yields for different quantities, combinations, and timings of nutrient applications. The model draws on standard nutrient response curves for tomatoes, which are based on aggregate amounts of N, P, and K application. However, standard agronomic models assume optimal timing of application. Since actual nutrient application in the field is often suboptimal, we worked to develop a model that adjusts standard yield curves with yield penalties for nutrient applications that depart from optimal timing of application. To do this, the agronomist generated yield predictions for a factorial combination of macronutrient quantities (Table A10) and timings (Table A11), drawing on standard nutrient response curves for tomatoes. He then modified these cells to incorporate yield penalties for poor timing of nutrient applications based on agronomic principles.

From this we trained a simple model, $\hat{g}(z)$, for returns to nutrient application over the whole space of nutrient applications and timings. We first calculated the parameters for a nutrient response curve by regressing log yields on raw nutrient levels. We then estimated yield penalties regressed log predicted yields on the percentage of each nutrient applied during each stage of plant growth (planting, vegetative growth, fruitset and flowering). We use the coefficients from the first model to generate a yield response curve, and then use the estimated penalties from the second equation to adjust for suboptimal timings. We repeated this for the demo "grey soil" and counterfactual "black soil". The agronomist predictions for fertilizer combinations, for the incentivized problems (prediction and adaptive experimentation) sets $\hat{f}(x) := \hat{g}(z(x))$, using the known macronutrient composition of fertilizers $z(x)$.

Nitrogen rate (kg/ha)	Phosphorus rate (kg/ha)	Potassium rate (kg/ha)	Black Soil Yield	Grey Soil Yield
0	0	0	4	3
0	30	0	5	4
0	30	40	10	8
0	30	80	16	12
0	30	160	21	16
0	30	240	17	13
20	0	0	7	6
20	30	0	12	9
20	30	40	13	10
20	30	80	16	12
20	30	160	24	18
20	30	240	26	20
40	0	0	17	13
40	30	0	18	14
40	30	40	24	18
40	30	80	26	20
40	30	160	29	22
40	30	240	28	21
80	0	0	23	18
80	30	0	26	20
80	30	40	28	21
80	30	80	30	23
80	30	160	32	24
80	30	240	30	23
120	0	0	26	20
120	30	0	26	20
120	30	40	28	22
120	30	80	29	22
120	30	160	30	23
120	30	240	28	22

Table A10: Agronomist predicted yields under different nutrient application rates. Here timing was held fixed: All the P applied at transplanting to the field. One quarter of each N and K was applied at transplanting then top-dressed with half of the full dose 3 weeks after transplanting. Lastly one quarter of the full dose was applied at flowering

Nitrogen (kg/ ha)			Potassium (kg /ha)			Manure (t / ha)	Phosphorus (k/gha)				Agronomist Predictions	
Planting	Vegetative	Flowering	Planting	Vegetative	Flowering		Planting	Vegetative	Flowering	Fruitset	Grey Yield	Black Yield
25	25	25	30	30	30	0	30	0	0	0	13	20
25	25	25	30	30	30	0	0	30	0	0	9	13
25	25	25	30	30	30	0	0	0	30	0	7	9
25	25	25	0	30	60	0	0	0	0	30	12	20
25	25	25	0	30	60	4	30	0	0	0	25	40
25	25	25	0	30	60	4	0	30	0	0	23	36
50	25	0	0	30	60	0	0	30	0	0	20	29
50	25	0	0	30	60	0	0	0	30	0	17	25
50	25	0	0	30	60	0	0	0	0	30	15	22
50	25	0	0	30	60	4	30	0	0	0	27	40
50	25	0	0	30	60	4	0	30	0	0	25	36
50	25	0	0	30	60	4	0	0	30	0	22	32
50	25	0	0	30	60	4	0	0	0	30	22	32
25	50	0	30	30	30	0	30	0	0	0	15	24
25	50	0	30	30	30	0	0	30	0	0	9	15
25	50	0	60	30	0	0	0	0	30	0	6	9
25	50	0	60	30	0	0	0	0	0	30	6	9
25	50	0	60	30	0	4	30	0	0	0	22	34
25	50	0	0	30	60	0	0	30	0	0	19	34
0	25	50	30	30	30	0	0	30	0	0	8	13
0	25	50	30	30	30	0	0	0	30	0	5	7
0	25	50	30	30	30	0	0	0	0	30	5	7
0	25	50	30	30	30	4	30	0	0	0	22	32
0	25	50	30	30	30	4	0	30	0	0	18	27
0	25	50	30	30	30	4	0	0	30	0	15	23
0	25	50	30	30	30	4	0	0	0	30	15	23
0	25	50	60	30	0	0	30	0	0	0	8	13
0	25	50	60	30	0	0	0	30	0	0	8	13
0	25	50	60	30	0	0	0	0	30	0	5	7
0	25	50	60	30	0	0	0	0	0	30	5	7
0	25	50	60	30	0	4	30	0	0	0	18	26
0	25	50	60	30	0	4	0	30	0	0	15	22
0	25	50	60	30	0	4	0	0	30	0	11	16
0	25	50	60	30	0	4	0	0	0	30	10	15
0	25	50	0	30	60	4	0	0	0	30	22	32
0	0	75	30	30	30	0	30	0	0	0	9	16
0	0	75	30	30	30	0	0	30	0	0	7	11
0	0	75	30	30	30	0	0	0	30	0	5	7
0	0	75	30	30	30	0	0	0	0	30	4	6
0	0	75	30	30	30	4	30	0	0	0	12	25
0	0	75	30	30	30	4	0	30	0	0	10	21
0	0	75	30	30	30	4	0	0	30	0	8	18
0	0	75	30	30	30	4	0	0	0	30	8	18
0	0	75	60	30	0	0	30	0	0	0	7	11
0	0	75	60	30	0	0	0	30	0	0	7	11
0	0	75	60	30	0	0	0	0	30	0	5	7
75	0	0	60	30	0	0	30	0	0	0	8	13
75	0	0	0	30	60	0	0	0	30	0	13	22
75	0	0	0	30	60	0	0	0	0	30	12	20
75	0	0	0	30	60	4	30	0	0	0	22	36
75	0	0	0	30	60	4	0	30	0	0	19	32
75	0	0	0	30	60	4	0	0	30	0	15	27
75	0	0	0	30	60	4	0	0	0	30	15	27

Table A11: A subset of Agronomist Yield Predictions for different nutrient application timings. Here we hold total N, P, K fixed and just change the timing. The agronomist predicted 169 timing combinations total