3867 Behavioral Dynamics Affecting Data Use in Cambodia

Saugato Datta, Bidisha Barooah, Christian Chihababo, Michael Steiner

Michael Steiner, USA --- steinermic@aol.com

Abstract

While Cambodia has made substantial strides in making high-quality agricultural data from regular Agricultural Surveys available to policymakers and researchers, the use of such data remains limited. This paper reports on the findings from a survey of potential users of agricultural data in Cambodia, which focused on potential behavioral barriers to data use in the Cambodian context. Through an online survey carried out in the summer of 2024, we identify several key features of the context in which decisions about agricultural data use are made, as well as the underlying decision-making dynamics and behavioral/cognitive biases that may affect decision-makers, that may be inhibiting the appropriate use of agricultural data. We find that while respondents have generally positive attitudes to data use, and are well-informed about potential sources of agricultural data, there continue to be gaps in their ability to access and use such data. Beyond these capacity issues, we also document that potential users of data in Cambodia are susceptible to behavioral phenomena like present bias or hyperbolic discounting and loss aversion, while also operating in a context of cognitive scarcity engendered by multiple competing deadlines and tasks. Combined with a relative paucity of public feedback about data use and few explicit guidelines about data use and evidence suggesting that potential users of data may be susceptible to social and environmental cues (including past examples) about work priorities, our results suggest that some simple interventions around public acknowledgement and feedback, among others, might help increase the utilization of agricultural data in Cambodia.

1. Background and Motivation

High-quality, relevant agricultural data is an important input into sound policy decisions about agriculture and allied sectors, and the scarcity of such data constitutes an important barrier to effective policy making, with downstream impacts on economic growth and poverty reduction. In part to address these barriers, a consortium of development agencies have come together as part of the "50x2030 Initiative", which "focuses on improving country data by developing a fit-for-purpose, integrated, and financially sustainable agricultural and rural survey program that fosters a culture of data use for decision-making to support agricultural sustainability and rural development, address food crises, and mitigate the impacts of climate change"[1].

The 50x2030 Initiative was developed to increase the capacity of participating countries to produce, analyze, interpret and use high-quality and timely agricultural survey data for evidence-informed decision and policy making. The Initiative's implementation is divided into three interdependent components – data production, methods and tools development, and data use.

Having a separate component for data use, the 50x2030 Initiative places high importance on promoting data use and identifying barriers to data use.

The structure and activities of the Initiative are underpinned by the recognition that while improving the quality and availability of agricultural data is critical, so too is its appropriate use by relevant stakeholders[2]. While use could be limited initially by the non-availability of data, a host of evidence from a variety of geographies underlines the fact that availability of data does not necessarily result in its use (ibid.). Beyond such 'structural' factors, behavioral science suggests that these decisions - such as whether or not to use and analyze primary data for a work product, such as a report, memo or proposal - depend on 'behavioral' factors including features of the decision-maker's context (such as social norms and signals, implicit directives, cognitive abundance/scarcity, etc.) as well as underlying features of the decision-maker's preferences - such as loss aversion (the tendency for losses to loom larger than equivalent gains) or the nature of their intertemporal preferences (preferences about choices over time), etc. - which may of course be exacerbated by features of the context[3]

While the need to address structural barriers to data use (such as data availability) is widely understood and the focus of ongoing efforts, including those being undertaken as part of the 50x2030 Initiative, less attention has been paid to the existence and possible relevance of potential 'behavioral' biases and factors.

Cambodia was the first country entering the 50x2030 Initiative. Prior to this engagement, researchers had limited access to micro-level data. The Cambodia Socio-Economic Surveys provided micro-level data, but only collected limited amounts of agricultural data on a country wide sample of urban and rural households, and data was not available by province. The availability of micro-level agricultural data for Cambodia changed beginning in 2021 with the release of data from the first (2019) round of the Cambodian Annual Agricultural Survey, which was carried out with the support of the 50x2030 Initiative. Annual Agricultural Surveys have now been conducted in Cambodia each year since 2019, with three rounds of data having been released to date, with two further rounds expected to be released by the end of 2024. Access to agricultural data, rated as "moderate" in the 2021 Data Use Survey carried out under 50x2030, has also improved. The gradual easing of structural barriers around data quantity, quality and availability makes it both important and an opportune time to explore any behavioural constraints around data use. This study therefore seeks to contribute to our understanding of the behavioral dynamics behind current patterns of data use in Cambodia as a step towards identifying additional interventions that could be employed to increase data use, in line with the 50x2030 Initiative's third pillar around promoting data use.
Through a structured online survey administered as described in the following section, we sought to understand the following broad questions:

Are potential users of Cambodian agricultural data ***motivated*** to use data? What are their ***beliefs and attitudes*** towards data use and data analysis?

Do potential users of Cambodian agricultural data have the ***capacity/ability*** to use such data as are available?

Is the ***professional, social and institutional context*** in which potential users of agricultural data in Cambodia work conducive to their using data? Are there indications that features of the context might be activating potential behavioral biases?

Finally, do potential users of data in Cambodia ***display patterns of decision-making consistent with some frequently observed features of decision-making*** which in turn can explain certain behavioral patterns relevant to data use decisions, such as procrastination/deferral and a reluctance to abandon familiar or well-established ways of operating?

2. Methodology

The survey conducted by the 50x2030 data use team was designed to gather insights into the Cambodian agricultural sector through an online questionnaire. The survey was open for three months from June to August 2024.

Participants in the survey were recruited through targeted outreach to Cambodian professionals in sectors and organizations that could potentially use agricultural data. They included a wide array of professionals from government agencies, research institutions, international organizations, and the private sector. The responses were gathered from individuals in diverse roles, including planning, statistics, policy advising, and project management. The typical respondent holds varying seniority levels, particularly among Chief Officers, Advisors, and Executive Directors, highlighting a strong presence of experienced professionals in agriculture, research, and development. This diversity ensured the survey captured a broad range of perspectives within the sector. Thirty-nine unique individuals responded to the survey, and their responses form the basis of the results discussed in this paper.

The survey covered several key themes, including the nature of participants' work, their experience with competing deadlines, and their approach to data analysis. Specific questions explored how respondents handle multiple tasks, manage unexpected changes, and incorporate data analysis into their projects. Additionally, the survey sought to understand the recognition of high-quality data work and the general decision-making processes of participants. This comprehensive approach aimed to reveal both the practical challenges and personal attitudes towards agricultural data use.

The average completion time for the survey was approximately 40 minutes, which allowed respondents to provide in-depth and thoughtful responses. This thorough approach to data collection ensured that the findings would be robust and representative of the sector's broad spectrum, contributing valuable insights to the research objectives.

We should caveat our findings by noting that while there is a robust literature on uncovering behavioral biases using surveys such as this one, we are naturally limited both by a small sample as well as the likelihood of biases introduced by phenomena such as courtesy bias. While

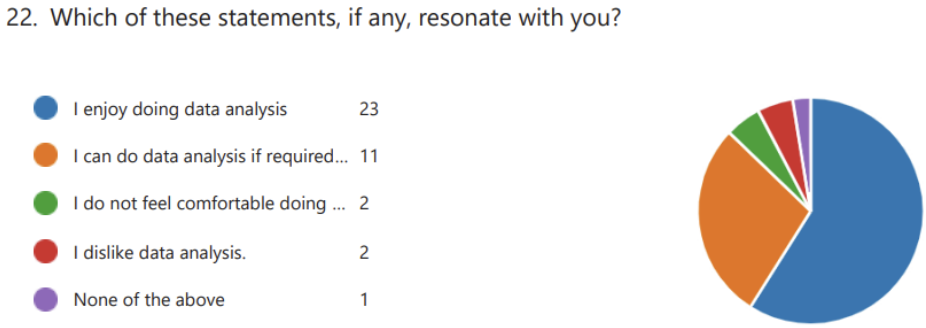
unavoidable in contexts such as this, this should nevertheless lead us to be cautious to interpret findings from a survey such as this. While we can speculate on the real-world implications of some of the phenomena we document, this should not be taken as definitive evidence for these effects, since we do not have the ability to observe real-world outcomes and must necessarily rely on respondents' (possibly not entirely accurate) accounts of how they would behave in specific situations.

3. Main Findings

Motivation and Attitudes: What are respondents' attitudes towards analytical tasks involving using data? What sorts of feelings does it evoke among them?

Result 1: Respondents generally report positive attitudes towards data analysis.

When asked directly, respondents generally expressed positive views of data analysis and their ability to carry out such analysis. 23/39 (59%) reported 'enjoying' data analysis, and only two respondents (5%) reported that they disliked it, with the remainder expressing neither enjoyment nor dislike along with various degrees of comfort with such tasks.



Result 2: Despite this, many respondents express some concern or trepidation about data analysis.

When asked to state the first two words that came to mind when contemplating a data analysis task, respondents' replies paint a more nuanced picture. While "interesting" was the most frequent response (21/39), almost as large a number of respondents (15/39) cited a word with a negative connotation (difficult, impossible, fear, scared, worried, annoying) and 6/39 cited 'time consuming'. Perhaps even more telling, of the 23 respondents who reported enjoying data analysis, 11 cited a word with a negative connotation (worried, difficult, impossible, fear, etc) as one of their top two associations with data analysis. This suggests that some people were perhaps hesitant to admit that they found data analysis stressful or difficult.

B: Ability/Ease/Comfort: Do potential users of agricultural data in Cambodia feel comfortable with their ability to carry out data analysis? Do they demonstrate knowledge of how to go about doing it?

Result 3: Respondents are broadly knowledgeable about where to obtain data needed for an analytical exercise, but do not necessarily only name the highest-quality sources

When asked where they would obtain the relevant data to carry out a hypothetical analytical task for a report, 31/39 (80%) people did correctly identify the Cambodian National Institute of Statistics (NIS) as one of the sources they could consult. However, 24 of these people also mentioned the data from the Ministry of Agriculture, Forestry and Fisheries (MAFF). Agricultural data collected by NIS is based on sound statistical methodology implemented in Cambodia under the 50x2030 Initiative, while data produced by MAFF has not had direct benefit from the 50x2030 Initiative, and thus, MAFF data is generally regarded as a lower-quality data source. This suggests that while a majority of respondents do cite the highest quality data source, most of them also cite a lower-quality source, though only 6 people cite only the MAFF data (ie the lower quality data source). Awareness about the availability and source of high-quality data has thus improved since the first Data Use Assessment carried out under the aegis of 50x2030 in 2021, when data awareness was very low. However, the ability of people to distinguish between the higher-quality source of data and other valid but lower-quality sources of data is still in need of improvement.

Result 4: Relatively few respondents can articulate valid next steps to obtain data needed for the hypothetical data analysis task, indicating some gaps in capacity to use data.

When asked for the next steps that they would need to take to obtain the data needed, a much smaller fraction (10/39) of respondents gave valid responses. This suggests that despite respondents' generally positive self-perceptions about their ability to use data, and indeed robust knowledge of *where* to obtain needed data, most respondents did not actually have the knowledge and skills necessary to obtain and analyze data, pointing to a major gap in capacity.

Result 5: Most respondents have never received explicit guidance on the use of data.

Only 17/39 (43.6%) respondents report ever having received guidance about the use of data as part of the instructions or brief for reports or other work products. Explicit guidance and guidelines might go some way towards remedying the lack of specific knowhow identified in the previous result. More broadly, the relative novelty of high-quality disaggregated micro-data on the agricultural sector means that technical guidance and training on its use will be key to improving data use, as reflected in the 50x2030 data use component on technical assistance.

C: Contextual Factors: Behavioral science has identified several ways in which a person's context may (often unconsciously) affect their decision-making, especially in situations where they are performing something new, unfamiliar or at least not routine. We therefore ask whether the professional, institutional and social context in which potential users of data find themselves is conducive to using data, or whether there are features of the context which may activate or exacerbate certain well-known behavioral biases that could affect data use?

Cognitive Scarcity: Data analysis is a time-consuming task, especially if potential users must locate, download and potentially clean data before analyzing it. Recent research in behavioral science focuses on the impacts of a scarcity of 'cognitive bandwidth' (or 'mental space'), finding that people's decision-making exhibits certain predictable patterns when people are busy or juggling multiple tasks[4]. Certain features of the context in which people operate - how busy they are, how many deadlines they are juggling, and so on - may thus help us understand why they may not always use data even when they could.

Result 6: Most users are operating in a context of cognitive and time scarcity.

A large fraction of respondents report dealing with multiple tasks and deadlines: 26/39 (66.7%) report that they are typically working on three or more work products at any given time, and 29/39 (74.4%) report that having competing deadlines at work is either "very common" or "somewhat common". In other words, the work environment for many potential users of agricultural data in Cambodia is characterized by time and cognitive scarcity. This finding suggests that there is a role for tools that make data analysis less time-consuming and more efficient, such as PowerBI.

7. Thinking of a typical week at work, are you typically working on one or more work products? If so, how many in a typical week?

8. How common would you say it is that you have more than one competing deadline at work - say two reports or other work products that are due at or around the same time.

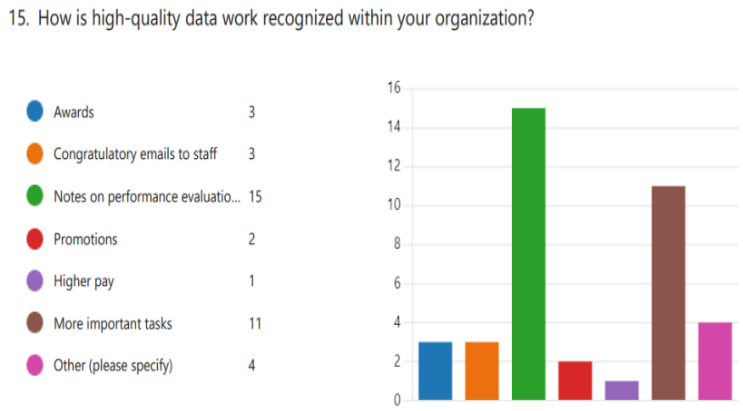
Result 7: The state of scarcity may affect the time and effort devoted to effortful tasks (such as data analysis).

Data analysis is a relatively time-consuming task and therefore one that may be neglected or deprioritized if a person is facing cognitive scarcity, which can lead to 'tunneling' (single-minded focus) on tasks or activities perceived to be particularly urgent[5] . We find that 15/39 (38.5%) respondents say that an unexpected addition to their workload would lead them to reduce the time and effort they would expend on their pre-existing task.

Norms and the Status Quo: At the same time, for most people who could use data, whether to use it is truly a decision, in that there are rarely explicit guidelines or guidance about whether and when to use data. Behavioral science finds that in such ambiguous situations, people often rely on precedent and are prone to sticking to what they have done before in a similar situation ("status quo bias"), or their perception of what the most acceptable course of action is ("social proof") when deciding what to do[6]. We should note that as summarized in Result 5, there is a dearth of explicit guidance about data use, making it more likely that people would rely upon subtle environmental cues to guide them in their work.

Result 8: Cues about data use from previous examples of similar tasks/products would affect many respondents' decision to use data or the time and effort devoted to data analysis.

15/39 respondents reported that finding out that previous approaches to a similar task did not employ data analysis would make them rethink their initial assessment of the need for data analysis, while another 9/39 said that while they would still carry out data analysis, it would now have lower priority. Thus, for a large fraction of the respondents (24/39 or 61%), the visibility of other similar work products that use data analysis may affect their decisions about whether to use data and/or how much importance to give it, suggesting that they are influenced by the perceived prevailing social norm around data use. We should note that the primary mode of feedback about good data analysis comes via performance reviews, which are typically private. There thus appears to be very little - if any - explicit *public* acknowledgment of quality data analysis and its contribution to organizational goals in respondents' workplaces, as well as little publicly advertised instances of data use having contributed to output quality, bolstering the likely role of perceptions of the norm or status quo, even if these are not accurate. More broadly, this finding points to the lack of incentives for data use.



Result 9: Respondents indicate a high degree of reliance on either existing models or guidance from others when deciding on how to approach a project.

When asked how they would react to being given a previous report to use as a model which seemed to take an inappropriate approach, most respondents (32/39) report either altering their approach to match the example, seeking out other examples, or seeking guidance from a supervisor. Only 7/39 respondents said they would carry on as they had initially planned, indicating a high degree of reliance on existing examples or the opinions/guidance of others.

D: **Behavioral Patterns in Decision-Making**: Decision-making has been documented to exhibit certain patterns that may help explain why people may sometimes defer or procrastinate about effortful tasks, or why they may be reluctant to depart from established patterns of action. We discuss our findings about such behavioral phenomena/biases below:

Result 10: About a quarter of respondents are "present-biased", suggesting that they place a disproportionate weight on the immediate, and may be prone to procrastination or deferral when it comes to effortful tasks.

Behavioral economists have documented a pattern of decision-making when it comes to choices over time, which is characterized by people putting an especially high weight on the present relative to any future period. This is called 'present bias' or 'hyperbolic discounting', and it explains many observed features of decision-making, such as repeatedly putting off (deferring) or procrastinating on tasks which require effort in the moment but lead to benefits, results or rewards in the future. Not all people exhibit this pattern of decision-making, but about a quarter to a third do in most studies.

We gave respondents a classic time preference task, where they were asked to choose between an immediate and a slightly higher deferred payoff, and then given the same options, but a year in

the future. A present-biased person will exhibit what is known as time inconsistency: they will express a willingness to wait a bit for a reward when all choices are in the future but will choose the immediate reward when the choice is imminent.

We found that 9/39 (23.1%) respondents exhibited a pattern of decision-making about delaying payoffs now and in the future consistent with present-biased preferences. This suggests that just under 25% of respondents display hyperbolic preferences - broadly in line with the literature. For reference, Ashraf Karlan and Yin (2006)[7] use a very similar question in the Philippines and find that 26% of their sample exhibits hyperbolic preferences.

We should note that case for interpretation of these numbers as being potentially relevant to decision-making around data use is bolstered by the finding that the main way people receive positive feedback for analytical work is via their performance reviews, which typically would happen only infrequently (usually once a year), so that 'rewards' (in the form of positive feedback) are delayed relative to the effort involved in data analysis.

Result 11: A substantial majority (70%) of respondents exhibit loss aversion.

Loss aversion refers to a well-documented pattern in decision-making, where losses cause a greater reduction in utility than equivalent gains increase utility. As a result, giving someone an amount of money and then taking it away leaves them worse off than if they had never received the sum of money in the first place. The reason is that receiving the money changes the person's baseline reference point, with respect to which they evaluate further changes.

To test whether respondents were loss averse, they were presented with a choice about two ways that they could receive monetary incentives for good work that are identical in expectation, but where one involves money being taken away for less-than-perfect performance ("loss frame") while the other does not ("gain frame"). See the figure below for the question and responses.

As predicted by prospect theory, respondents overwhelmingly displayed loss aversion. Of the 34 people who expressed a clear preference, only 10 chose the loss-framed option while 24/34 (70%) chose the gain-framed option, thus exhibiting loss aversion.

19. Imagine that your organization decides that people working on difficult projects will be given a small monetary bonus ($100) for each completed project up to four projects a quarter.

You have two choices for how to receive these bonuses: Either

Option A: You can take $400 at the beginning of the quarter and return $100 for each project you do not complete. So, if you complete two projects this quarter, you have to return $200 to the company. If you complete three projects, you'd return $100, and so on.

Option B: You get $100 for each project you complete, up to a total of up to $400. So, if you complete three projects, you get $300, and so on.

Assume you can't earn any interest, and there is no inflation. Which option would you prefer?

20. Have you ever begun to tackle a task and had realised you would need more time than you had estimated to complete the task? How did this affect your work on the task?

Loss aversion is a possible explanation for people's reluctance to deviate from established ways of approaching activities or tasks, known as the status quo bias[8]. To the extent that data analysis may not always be an established part of how certain tasks are approached (possibly in part because relevant data was not available in the past), the prevalence of loss aversion may explain why relatively few people start using data when it is made available.

Result 12: Virtually all respondents report being prone to the planning fallacy, and nearly 75% report that it affects their approach to their work.

Planning fallacy refers to people's tendency to systematically underestimate how long tasks, including ones they have carried out earlier, will take. Prior evidence suggests that people are often prone to the planning fallacy, and that this can lead them to abandon such tasks. When asked about whether they have had to revise their estimates of the time a task would take, and whether this caused them to adjust their approach to the work at hand, around 50% of respondents (21/39) answer that they have faced this situation and that this altered their approach to the task at hand, suggesting the prevalence and impact of planning fallacy. Another 8 respondents report that they have faced this situation, and that it affected other tasks/aspects of their work, and an equal number (9/39) say that they faced the situation, but it did not affect their

work. Thus, virtually all (38/39) respondents report facing planning fallacy, and a majority of these (29/39) say that it affected their work via one task or another. Most respondents in this case would appear to be susceptible to this phenomenon, which could lead to data analysis tasks being abandoned midway.

4. Conclusion and Recommendations

Within the limitations imposed by a short survey, and acknowledging the potential role of say-do gaps, this study has uncovered some interesting features of the decision-making of potential users of agricultural data in Cambodia.

First, the study finds encouraging evidence on potential users of data being aware of key sources of agricultural data. This is a marked change from the findings of the first Cambodian Data Use Survey, carried out by 50x2030 in 2021, where "poor information on sources of data and analysis" was the most common reason cited for low access to data, with 69% of respondents citing this reason. Awareness was also among the top three constraints on the data ecosystem in all but one of the seven sources of data studied, and was the second-most prevalent constraint overall, i.e., across all data sources. In contrast, we find that 87% of respondents correctly identify valid sources of agricultural data, suggesting that there has been a big improvement in awareness about data sources since 2021. As noted earlier, three rounds of data from Cambodia's Annual Agricultural Survey have now been released, with two more rounds expected to be released by the end of 2024, filling a major gap in data availability and access. The findings here thus find that the 50x2030 initiative has succeeded to a great extent in improving the basic data infrastructure in Cambodia's agricultural sector, and that awareness has improved in line with greater availability of data.

However, we also find results which suggest that interventions going beyond making data available and accessible may be needed to increase its use.

First, very few respondents report receiving explicit guidance about when and how to use data. Read in combination with the finding that only about a quarter of respondents offer valid responses about how they would go about obtaining and analyzing data, this suggests that there is a need for explicit guidance about when to use data and what kinds of data analysis to carry out, as well as training potential users of data about available sources, how to access or request them, and the nuts and bolts of obtaining and analyzing agricultural data. This might need to be supplemented with more training on how to access, analyze and interpret data. Currently, IFAD is providing some training on a limited scale, and mainly to Ministry of Agriculture staff. To make a big impact in a country, training needs to be provided to many sectors --- including NGOs, academic institutions, and media. Other participants in the 50x2030 Initiative can also help provide trainings across a wide spectrum of stakeholders.

Explicit guidance about when and how to use data may also bolster its use in a variety of ways. First, it would make the social desirability of data use more visible, and over time, would allow data producers to highlight the dynamic social norm of the increasing frequency of data use. Secondly, such guidelines would help to make data use a priority, even when people are facing numerous other tasks and deadlines, as survey respondents indicate is usually the case.

Thirdly, the results here suggest that there is little public praise or acknowledgement of good data use. Providing those who do use data with public acknowledgment or praise would help to make data use salient and act as an implicit recommendation, while also reducing the gap in time between data use and feedback on the same. This would make it less likely that people would defer or procrastinate about data use, a tendency that is heightened when effortful actions are carried out well in advance of any benefits or positive feedback that may arise from it.

Finally, we find that potential users of data routinely underestimate how long tasks will take. Given that evidence from the behavioral sciences suggests that this can lead to discouragement and abandonment of tasks, it may be helpful to provide users with realistic estimates of the time and effort required for certain kinds of common data analytical tasks, helping to set realistic expectations and reduce the likelihood that users will get discouraged and give up on data tasks, especially when such activities are relatively unfamiliar to them. An example of this kind of intervention is forms or surveys that give people an estimate of how long they will take the typical person to complete them.

Overall, our findings suggest that agricultural data use is not explicitly called out or highlighted in public ways, but that doing so would be helpful in reducing the effect of certain behavioral biases, such as hyperbolic discounting or status quo bias. Past research suggests that simply highlighting desirable practices helps people see them as socially desirable and can increase the uptake of such practices.

Given that we find that the use of agricultural data in Cambodia is likely inhibited by a number of behavioral patterns and biases, such public acknowledgment – bolstered, if possible, by measures such as prizes or awards for good quality data work – would potentially go a considerable way in helping unlock the benefits of the investment in higher quality, more frequent and reliable agricultural data (an area in which Cambodia has made substantial strides in recent years by fielding regular Agricultural Data Surveys) by encouraging its use in policy analysis and formulation.

References

[1] 50x2030 Initiative Website (https://www.50x2030.org/about). Accessed September 2024.
[2] "Data & Policy: A new venue to study and explore policy–data interaction". Verhulst, SG, Engin, Z and John Crowcroft. 2019. Data & Policy, Volume 1 , 2019. DOI: https://doi.org/10.1017/dap.2019
[3] Behavioral Insights on the Use of Evidence for Local Policy-Making in South Africa. by Laura Cojocaru, Saugato Datta, Abigail Sellman, and Cecilia Shang. ideas42 White Paper, November 2021
[4] Mullainathan, S., & Shafir, E. (2013). Scarcity: Why having too little means so much. New York: Henry Holt and Company.
[5] Shah, A. K., Mullainathan, S., & Shafir, E. (2012). Some consequences of having too little. Science, 338(6107), 682–685. https://doi.org/10.1126/science.1222426.
[6] Cialdini, R. B. (1994). Principles and techniques of social influence. In A. Tesser (Ed.),
Advanced social psychology. New York: McGraw-Hill.
[7] "Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines"Nava Ashraf, Dean Karlan, Wesley YinThe Quarterly Journal of Economics, Volume 121, Issue 2, May 2006, Pages 635–672, https://doi.org/10.1162/qjec.2006.121.2.635
[8] Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias". Daniel Kahneman, Jack L. Knetsch, Richard H. Thaler Journal of Economic Perspectives vol. 5, no. 1, Winter 1991 (pp. 193–206)