Does citizen monitoring improve infrastructure quality?

A statistical comparison of monitored and unmonitored projects in Kwale County, Kenya

This document provides analysis of data collected from 283 infrastructure projects in Kwale County, Kenya, which were implemented between 2018 and 2022 under the departments of medical and public health, water services, or education.

Many of these projects were monitored during construction by local citizens participating in the VOICE initiative supported by KYGC and Integrity Action. Through this initiative, citizens acted as community monitors to raise problems with projects as they occurred and to work with those responsible for achieving solutions.

The analysis presented here was carried out by Daniel Burwood, Evidence & Impact Manager at Integrity Action, to examine whether any differences existed between the projects that had been monitored and those that had not.

A summary of findings is as follows:

i. There is a strong, positive, and statistically significant relationship between a project's condition and whether it was monitored during implementation.

ii. This relationship exists for projects that were monitored only by their project management committees (PMCs), though is strongest when community members are involved – either without PMCs, or when working with them in collaboration.

iii. This relationship existed even when monitoring began in a year after the start of project implementation.

iv. This relationship does not by itself prove causality, i.e. it is not possible to conclusively say from this data that monitoring leads to better project conditions. A dedicated study would be needed to robustly test any causal claims, which was not within the scope of this exercise.

v. The distribution of monitoring was not random across the locations covered by the study, i.e. some wards had a larger percentage of their projects monitored than others. A strong, positive, and statistically significant relationship also exists between a project's condition and its location. This is consistent with monitoring leading to better conditions, though is also consistent with alternative explanations such as certain local officials being more open to monitoring and better able to deliver good condition projects.

vi. Projects reported as complete and in good condition are statistically more likely to be in use by their communities, as expected. However, there is no apparent direct link between a project being in use and whether it was monitored.

vii. There is indicative evidence of monitoring being linked to other desirable behaviours or outcomes, beyond those projects reported as completed and in good condition. One such link could be that monitored projects are more likely to be stalled rather than handed over in poor condition, potentially keeping open opportunities for improvement if not quite fixing problems. Another is that unmonitored projects may be more likely to be vandalised, become unusable, or otherwise fall into disrepair after completion. Both links are backed by evidence, but with low sample sizes they would need further exploration through another study.

Contents

1. Background

This analysis is based on data collected between 19 September and 13 October 2022 in Kwale County, Kenya. Data was collected by 15 research assistants trained and supported by KYGC, each of whom spent three or four days visiting infrastructure projects in their allocated areas and recording the status of each one using the KoboToolbox form located in the annex of this document.

Projects were visited across six wards in which KYGC and Integrity Action had implemented their VOICE programme since 2018. This programme involved local community members taking the lead on monitoring delivery of county-funded projects and reporting any problems to the responsible authorities. The primary purpose of data collection was to see whether such monitoring activities were associated with any difference to the quality of completed projects.

Data collection began with a one-day planning meeting involving the VOICE team from KYGC and all 15 research assistants. It was decided that each assistant would target visiting ten projects a day, with the exercise expected to involve four days of data collection from each assistant. Four assistants in fact completed their observations in three days, whilst the others took the expected time but reached fewer projects than planned.

Challenges to reaching the targeted number of projects were largely due to the size of the areas being covered, with long distances between projects being especially difficult to cover due to the rainy season. However, research assistants received notable cooperation and support from project duty-bearers and local communities, and a total of 283 unique projects were visited through the exercise. This remains a very robust sample for analysis.

2. The project sample

Data was collected from across Kinango and Matuga sub-counties, covering the six wards of Kasemeni, Kinango, Kubo South, Mkongani, Tiwi, and Waa-Ngombeni. The status of each project was chosen from a closed set of options:

1. Complete and in good condition
2. Complete and in poor condition
3. Stalled
4. Ongoing

Overall, the majority of infrastructure projects visited were recorded as "complete and in good condition"; this status was attributed to 200 of the 283 visited projects. A further 51 were logged as "complete and in poor condition", 23 were "ongoing", and 9 were "stalled". Since it is impossible to say from the data whether an ongoing project is going well or poorly, these 23 have been removed from the analysis.

The assistants' choice of status was subjective and based on observation and judgment, rather than a fixed checklist or set of metrics, though the large sample size mitigates against inconsistency. Both "good" and "poor" condition are broad categories, with the latter theoretically covering the range from faded paintwork to severe structural damage. Free-text descriptions of the project were collected to help with this; these were provided for over 85% of cases and provide a sense of the interpretation being used by research assistants:

- Approximately 60% of projects recorded as in poor condition received comments that explicitly mention needing to replace broken, faulty, or missing infrastructure. Rusty taps and broken window latches were common features, while other projects suffered from leaking pipes, exposed wiring, rough or broken floors, stolen doors, cracked water tanks, or equipment such as pumps or solar panels not working. Some also noted use of substandard materials, such as incorrect timber or stands that are too weak to hold the weight of their water tanks.

By contrast, only 5% of "good condition" projects received comments about necessary repairs. This indicates that construction quality and maintenance played a key role in assistants' perception of condition.

- Some projects received comments about accessibility issues, especially for people with disabilities – often due to ramps being steep or absent. These issues affected about 10% of poor condition projects and only 2% of those in good condition, suggesting that accessibility was a factor in assistants' ratings.
- Projects reported to be in good condition often received praise for being well planned and organised, in good use, and being appreciated by the community. These show that being delivered as expected and functioning as intended were key factors for being seen as in good condition. That said, not all good condition projects were yet ready for use, as discussed in section 7 of this document.

Data was also collected on whether the project implementation had been monitored. Such monitoring could be done by community monitors (as in VOICE), by the project management committees (PMCs)[1], by a combination of both, or by other actors. The frequency of each type of monitoring is shown in the following table:

Monitored by	Complete, good condition	Complete, poor condition	Stalled	TOTAL
Community monitors *(only)*	50	5	3	58
PMCs *(only)*	53	12	1	66
Community monitors *and* **PMCs**	55	5	4	64
Another organisation	1	-	-	1
No monitoring	41	29	1	71
TOTAL	200	51	9	260

Since it is unclear what monitoring approach was taken by the one recorded as "another organisation", this project has also been removed from the analysis[2].

The remaining 259 projects were spread across the following locations:

[1] Project management committees are a structure mandated by Kenya's constitution to play a local oversight role in government-funded projects.

[2] Since the removed project was "complete and in good condition", its inclusion would if anything strengthen the case for monitoring.

Sub-county	Ward	Village unit	# monitored	# not monitored	Total
Kinango	**Kasemeni**	Bofu	21	-	21
		Minyenzeni	9	16	25
	Kinango	Gandini	12	10	22
		Kinango	20	4	24
Matuga	**Kubo South**	Majimboni	3	19	22
		Mangawani	13	-	13
		Mwaluvanga	6	3	9
		Shimba Hills	3	6	9
	Mkongani	Kizibe	14	-	14
		Mkomba	9	-	9
		Mlafyeni	18	-	18
		Tiribe	14	-	14
	Tiwi	Mkoyo	5	11	16
		Simkumbe	26	2	28
	Waa-Ngombeni	Kiteje	3	-	3
		Ngombeni	12	-	12

The research assistants also recorded the year in which implementation had begun for each project, as below[3]:

Pre-2018	2018	2019	2020	2021	2022
53	52	60	42	39	12

The department for each project was recorded as either education, medical and public health, or water services. The names of the projects were also provided, and from these it has been possible to deduce the more specific information on project type shown in the table on the following page. This information is provided to aid understanding of the sample, but the additional detail has not been used in further analysis. The large number of early childhood development centres reflects Kwale County government's priority during this period to enhance access to education.

Finally, to aid in verification, a photo was also uploaded of each project, and the GPS coordinates of each site captured. These do not feature in this document.

[3] Note these figures add to 258, as there was one project for which the start date has been left blank. This project has been kept in the sample where it does not affect the analysis, although for regressions it has had to be removed.

Project type	# monitored	# not monitored	TOTAL
ECDE	81	28	109
Vocational Training Centre	4	4	8
Education - other	3	-	3
TOTAL EDUCATION	**88**	**32**	**120**
Dispensary	15	6	21
Maternity wing	8	2	10
Staff quarters	6	2	8
Health - other	13	-	13
TOTAL HEALTH	**42**	**10**	**52**
Borehole	13	4	17
Dam	5	2	7
Pipeline	11	3	14
Water kiosk	9	-	9
Water tank	2	4	6
Water - other	18	16	34
TOTAL WATER	**58**	**29**	**87**

3. Limitations of the study

In conducting this analysis, the following limitations have been identified and borne in mind. These should be further considered by those interpreting or acting on these findings.

1. The reliance on a binary distinction between broad categories of "good" versus "poor", when considering the condition of completed projects. This is discussed in more depth in the previous section. While the research assistants' general approach to deciding on these seems clear, there will inevitably have been some projects that were on the borderline and that different assistants may have rated differently. Moreover, the data does not allow for further divisions within categories, for example distinguishing between a "poor condition" project that needs new doors and windows fitting versus one that needs complete demolition.

   The large sample size is expected to mitigate against much of the inconsistency in categorisation, but the lack of more nuanced or defined information does mean that less analysis is possible than if more detail on condition had been collected.

2. Details collected on the nature of the projects were also limited, and it is not clear the extent to which projects are comparable. While the three departments are clear, and so there is no real risk as above of projects sitting on the borderline, each department is again broad and there are substantial differences between a borehole, a water kiosk, a dam, a pipeline, and so on.

This limitation does not negate the analysis presented in this document, but does mean that it likely presents a less clear picture than if information had been available on, say, the budgets (planned and actual), contractors, exact timelines, composition of PMCs, processes of public participation, and so on. Specifically, without this information, drawing any conclusions about causality would be done at considerable risk.

3. The category of "monitored" is again broad, and may cover different forms of activity. While data is present (as shown in the section above) on whether projects were monitored by community members, PMCs, or a combination of the two, there will be differences within each category on how this was carried out. For example, KYGC & Integrity Action's VOICE programme went through three phases during the period from 2018-22, with some notable differences in approach between each phase[4]. It is also likely that there were differences in ability between individual monitors, or between different PMCs.

   This means that, while the comparison in this analysis of "monitored" versus "unmonitored" remains valid, it is not possible to present as clear a picture of what model or features of monitoring are most important to the relationship.

4. Are monitored projects in better condition than others?

This is the primary question of the analysis, and to answer it a chi-square test of independence has been used. This test provides a measure of whether the difference between an observation and an expectation is statistically significant, or due merely to random chance.

For this analysis, projects recorded as "complete and in poor condition" have been combined with those that were "stalled". These projects have then been contrasted with those that were "complete and in good condition". The contingency table is thus:

	Good condition	Poor or stalled	TOTAL
Monitored			188
Not monitored			71
TOTAL	199	60	259

If there were no relationship between monitoring and the condition of a project, then the expected distribution of "good" and "poor or stalled" projects would be split according to an even weighting. For example, since 199/259 (77%) of the projects were in good condition, we would expect 77% of the 188 monitored projects to be in good condition (144) and the remaining 23% to be poor or stalled.

[4] One of the biggest would be the adoption at the end of 2019 of Integrity Action's upgraded DevelopmentCheck tool. This provided monitors in this new "VOICE2" period with a checklist of potential problems to use for each project, as well as the ability to capture feedback from their local communities.

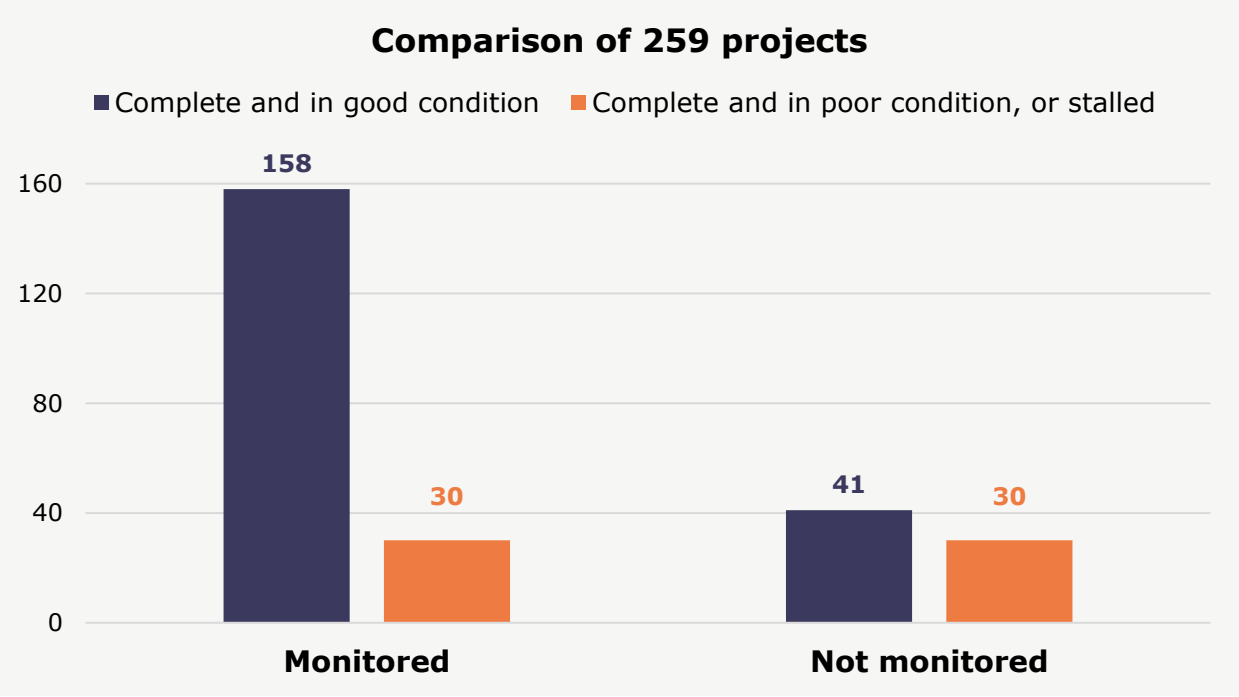
However, the data collected by the research assistants provides a different picture. The table below shows the ***expected value*** for each case (calculated as above), as well as the **observed value** that was actually recorded:

	Good condition	Poor or stalled	TOTAL
Monitored	**158** > *144*	**30** < *44*	188
Not monitored	**41** < *55*	**30** > *16*	71
TOTAL	199	60	259

This result shows that the number of monitored projects in good condition is higher than we would expect if there were no relationship, and so a positive relationship is possible. The chi-square (χ^2) test then assesses the likelihood of this difference being random. First, the χ^2 value is calculated by comparing the observations to the expectations, and this is then compared to the χ^2 critical value for the appropriate degrees of freedom and desired significance level.

In this case, χ^2 is 20.02, which is much greater than the critical value 6.63 at significance level $\alpha = 0.01$, and since the p-value is 0.0002 we can conclude that **there is a very strong and statistically significant relationship between a project's condition and whether it was monitored**[5].

The difference is also clearly illustrated in the following graphs:



[5] The significance level α is the probability of mistakenly identifying a relationship between two unrelated factors, so a low α provides higher confidence. Standard practice is to consider a relationship significant at $\alpha = 0.05$ or below.
The p-value measures the probability of randomly getting a higher value for χ2 than was observed, in a case where there is in fact no relationship. So a low p-value is again better, and indeed p needs to be lower than α for the test to hold.

5. Does it matter who monitored?

The same test as above can be used to check whether the relationship exists only for certain kinds of monitor. In the tables below, the projects monitored by each group of actors have been compared to the unmonitored projects. These again show the ***expected value*** for each case (if there were no relationship), and the **observed value** (as recorded by research assistants):

	Good condition	Poor or stalled	TOTAL
Community monitors *(only)*	**50** > ***41***	**8** < ***17***	58
Not monitored	**41** < ***50***	**30** > ***21***	71
TOTAL	91	38	129

	Good condition	Poor or stalled	TOTAL
PMCs *(only)*	**53** > ***45***	**13** < ***21***	66
Not monitored	**41** < ***49***	**30** > ***22***	71
TOTAL	94	43	137

	Good condition	Poor or stalled	TOTAL
Community monitors *and* **PMCs** *(together)*	**55** > ***46***	**9** < ***18***	64
Not monitored	**41** < ***50***	**30** > ***21***	71
TOTAL	96	39	135

In each case, the observed number of projects in good condition is higher than would be expected for the monitored projects if there were no relationship. The χ^2 tests are as follows:

Community monitors only

$\chi^2 = 12.44 >$ critical value 6.63 at $\alpha = 0.01$, $p = 0.006$

There is a very strong and statistically significant relationship between a project's condition and whether it was monitored by community monitors acting alone *(compared to no monitoring).*

PMCs only

$\chi^2 = 8.08 >$ critical value 3.84 at $\alpha = 0.05$, $p = 0.044$

There is a strong and statistically significant relationship between a project's condition and whether it was monitored by a PMC acting alone *(compared to no monitoring).*

Community monitors and PMCs working together

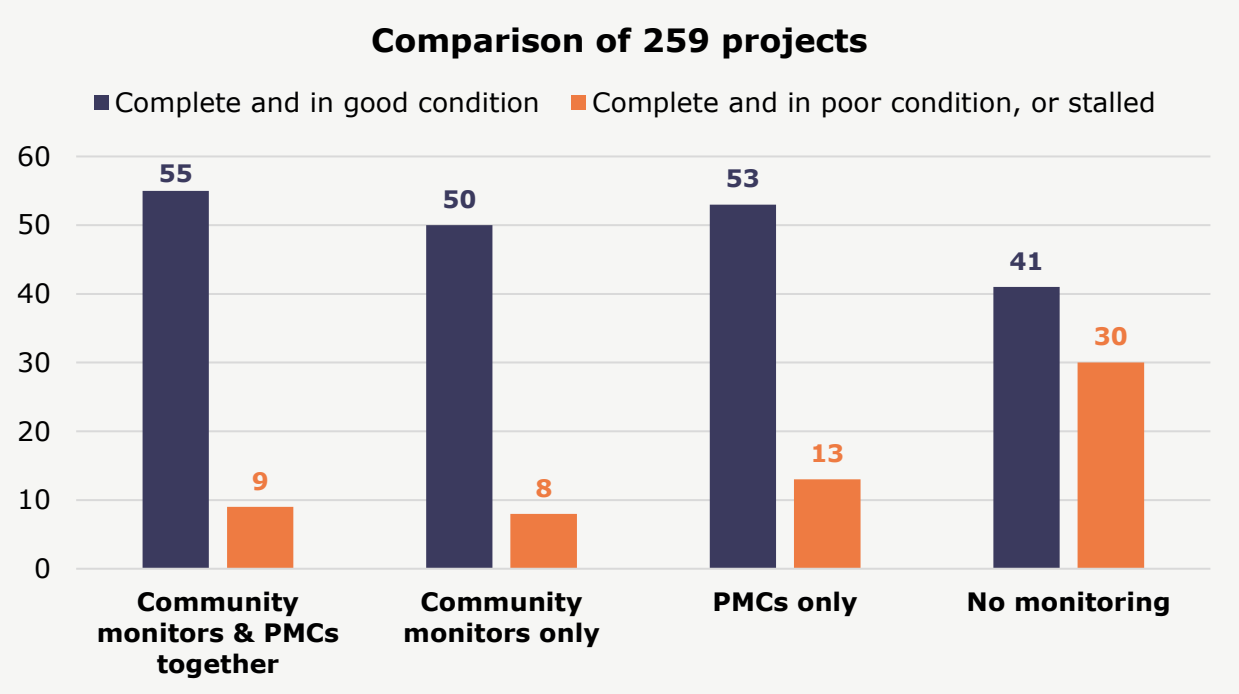
$\chi^2 = 13.02 >$ critical value 6.63 at $\alpha = 0.01$, $p = 0.005$

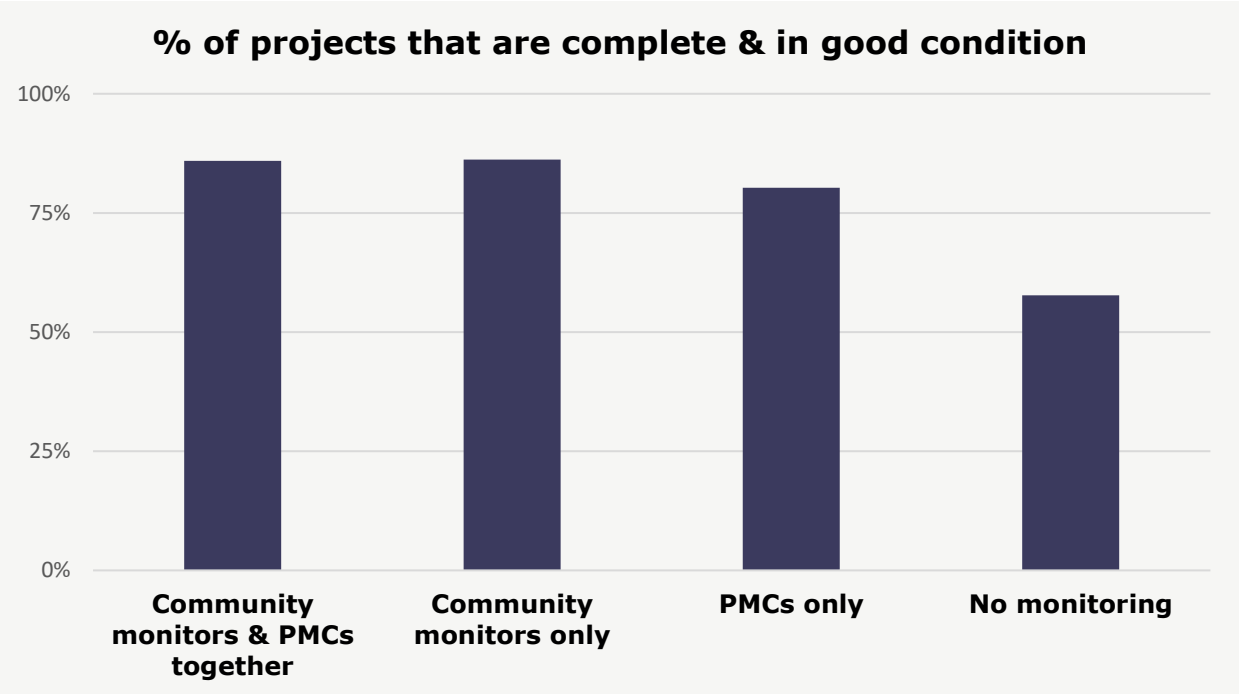
There is a very strong and statistically significant relationship between a project's condition and whether it was monitored by community monitors and PMCs working together *(compared to no monitoring).*

The χ^2 values show that the differences between expectations and observations are highest when community monitors are involved (either with or without PMCs). The p-value for the PMC-only case is also higher, meaning that the test is only valid for a higher value of $\alpha = 0.05$, suggesting a slightly weaker relationship[6]. This suggests that the connection between monitoring and the project condition is likely not as strong when PMCs act by themselves, and could indicate the specific value of community monitors.

These relationships are again illustrated in the graphs on the following page.

[6] The $\chi 2$ value of 8.08 is still greater than the critical value 6.63 at $\alpha = 0.01$, but in this case $p > 0.01$ and so that level of α cannot be used. However, $\alpha = 0.05$ is still widely accepted as a significant result.





6. Does monitoring lead to better projects?

The analysis above shows that there is a positive relationship between monitoring and project condition, but it does not prove that monitoring is the reason for better conditions. There may be other causes that affect both factors: for example, it could be that health projects tend to be delivered in better condition than other projects and are also more likely to be monitored.

Strict causality is difficult to prove conclusively, and would require a purpose-built data collection exercise considering all possible influences on a project's condition (e.g. details of the contractor, the budget, the quality of public participation at the planning stage, and so on). However, we can test some of the possibilities through the data we have; namely the department, location, and start date of each project.

As a first step, each pair of factors has been checked for any relationships. In most cases this has been done using the same chi-square test of independence as above, although Fisher's exact test has also been used where necessary[7]. The results are shown in the table below.

	Location	Department	Start date	Condition
Was project monitored?	Related	No relationship found	No relationship found	Related
Condition	Related	No relationship found	No relationship found	
Start date	No relationship found	No relationship found		
Department	No relationship found			

These results show that most of the factors are unrelated, e.g. there is no connection between the start date of a project and whether it is an education or health project. The relationship between a project's condition and whether it was monitored has already been described, but there are two more relationships that have been identified.

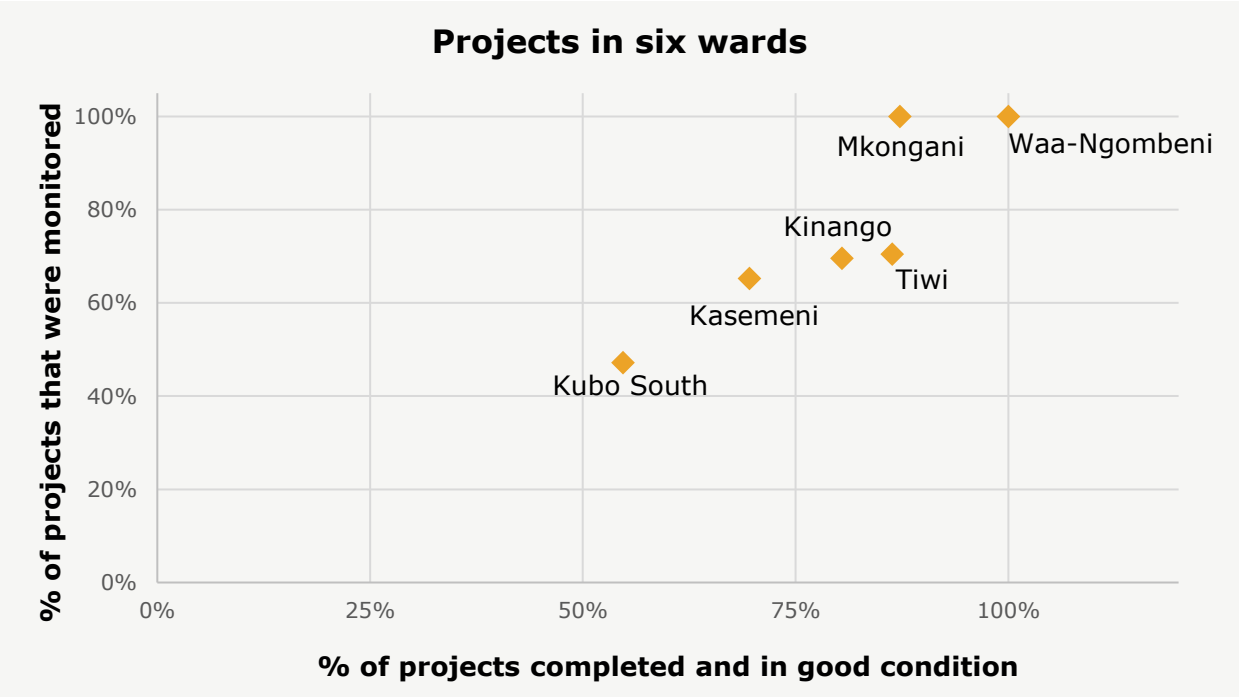
What the two relationships in the "location" column show are that the projects reported in good condition were not randomly distributed among the study's wards or villages[8], and neither were the projects that were monitored. The distributions are in fact as follows:

Ward	% projects that were complete and in good condition	% projects that were monitored
Kubo South	54.7%	47.2%
Kasemeni	69.6%	65.2%
Kinango	80.4%	69.6%
Tiwi	86.4%	70.5%
Mkongani	87.3%	100.0%
Waa-Ngombeni	100.0%	100.0%

[7] The χ^2 test relies on all values in the observed/expected table to be greater than 5, otherwise it is not strictly valid. Fisher's exact test is a statistical method that similarly compares two factors and can handle lower numbers.

[8] Comparisons of "Location" with other factors have been repeated using both the list of six wards and the list of 16 village units; the same relationships were found in each case.

Here we can clearly see that some wards were better than others for having projects in good condition, and also that some wards had much higher rates of project monitoring. Reinforcing the relationship that we already knew about from section 3, we can also see that the locations where more projects were monitored are also those where more projects are completed and in good condition. This is illustrated below:



There are several possible explanations for these findings, noting that geographic distribution of monitoring was never intended to be random:

1. Some wards and villages may have seen a higher proportion of projects monitored due to a circumstantial factor such as where monitors themselves are located. Indeed, the size of some wards meant that covering them entirely was not possible through VOICE. According to this explanation, monitoring has then led to better projects in these areas.

2. Some wards and villages may have seen a higher proportion of their projects being monitored due to an unrecorded causal factor such as having more co-operative officials or contractors, or more engaged local communities. This same causal factor would then have also delivered better projects in these areas.

3. Some unintended bias may have been introduced to the sample through the data collection process. For example, it is possible that in Waa-Ngombeni or Mkongani the research assistants visited only monitored projects due to being more aware of them, while perhaps the research assistants for Kubo South were able to draw on a more comprehensive list of recent projects.

Evidence gathered through, or on behalf of, the VOICE programme offers a compelling argument for explanation #1 – for example, the external evaluation of the VOICE2

phase of the programme, and the external evaluation of Integrity Action's 2020-24 grant from Sida (which includes a case study on VOICE). The data analysed in this document certainly supports this other evidence, but cannot by itself rule out the other explanations.

All the statistical tests presented above have compared just two factors at a time. More advanced techniques can be used to take into account how factors may be interrelated. These include logistic regressions, which check relationships between a dependent variable and each defined independent variable, while controlling for the other independent variables. In this analysis, the following logistic regressions have been run[9]:

i. Dependent variable: project condition. Independent variables: whether the project was monitored, village units, departments, start date

ii. Dependent variable: project condition. Independent variables: who monitored the project (if anyone), village units, departments, start date

iii. Dependent variable: project condition. Independent variables: whether the project was monitored by community monitors (with or without PMCs), village units, departments, start date.

Due to the relationship between location and monitoring status, interaction variables were added to each regression in an attempt to compensate for multicollinearity. However, it is likely this relationship still had a notable effect on the results.

Full tables of results have not been included in this document, but in summary:

- Projects that began implementation in 2022 were identified as less likely to have been completed in good condition than projects that began earlier. In fact, of the 12 projects that began in 2022, six are complete and in good condition, one is complete and in poor condition, and five are stalled (which is 5/9 or 56% of all stalled projects). This intuitively makes sense, as projects take time to complete.
- Education projects tended to be in better condition than other departments once other factors were controlled for. This may reflect the county government's priority in this area. In turn, health infrastructure projects tended to be better than water.

[9] For further clarity, the differences between these tests are:

Regression (i) contained a variable that =1 if the project was monitored and =0 otherwise;

Regression (ii) contained columns for each of the three types of monitors, which each =1 if monitored by this group and =0 otherwise; and

Regression (iii) contained a variable that =1 if the project was monitored by community monitors (with or without PMCs) and =0 if the project was not monitored or was monitored only by PMCs.

- The first regression failed to identify the relationships between condition and monitoring status or location. However, this does not undermine the findings from earlier sections; rather it highlights the difficulty in the regression controlling for changes in independent variables that are strongly related to each other.
- The second regression suggested a positive, though weakly-significant, relationship between project condition and monitoring by community monitors and PMCs working together (compared to no monitoring).
- The third regression suggested a positive, though weakly-significant, relationship between project condition and monitoring by community monitors (with or without PMCs), compared to no monitoring or monitoring only by PMCs.

These findings are consistent with analysis from earlier sections. The identified relationships appear weaker, although it is impossible to tell how much this is a result of the relationship between monitoring and location.

One final test we can do is to narrow the analysis to just those locations that had a mix of monitored and unmonitored projects. We will also remove projects that began before 2018 (when monitoring began under VOICE) in order to improve the comparison.

Focusing only on projects where implementation began between 2018-22 leaves us with 205 projects across all 16 village units. In 11 of these villages, all or nearly all of these projects were monitored; while in one village (Majimboni) fewer than 10% were. These are removed and this test then considers only the remaining four village units, where no more than 75% of projects were either monitored or unmonitored. This reduced sample is as follows:

Village unit	Monitored projects (started 2018-22)	Unmonitored projects (started 2018-22)	TOTAL
Minyenzeni	9	9	18
Mkoyo	4	9	13
Mwaluvanga	6	2	8
Shimba Hills	2	5	7
TOTAL	21	25	46

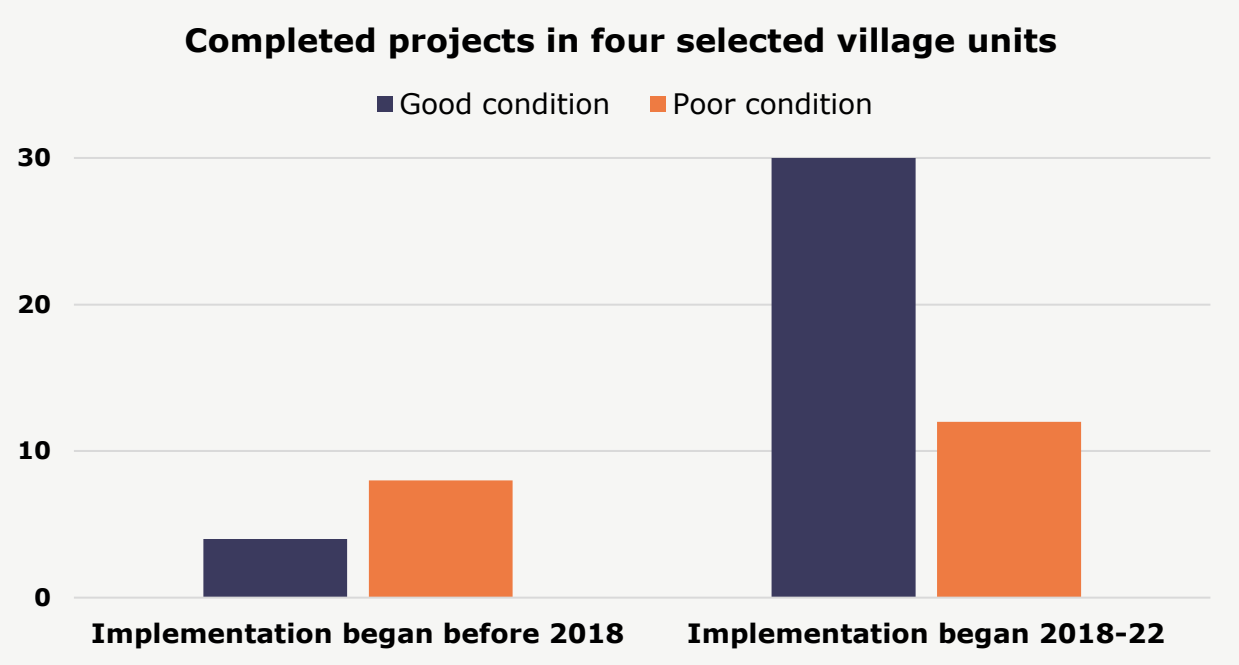
The next table provides details of these 46 projects:

	Good condition	Poor condition	Stalled	TOTAL
Monitored	14	4	3	21
Not monitored	16	8	1	25
TOTAL	30	12	4	46

In these four village units, approximately two-thirds of projects implemented since 2018 are complete and in good condition. This is true for both monitored (66.7%) and

unmonitored projects (64%), making a relationship difficult to identify. However, it's worth noting the change over time across these four village units.

Between these four village units, there were 12 projects where implementation had begun before 2018 (two that were monitored later, and ten that remained unmonitored), and as of 2022 only a third of these 12 were complete and in good condition. The improvement is shown in the graph below:



This suggests that citizen monitoring in these locations has coincided with a notable improvement in average project quality, although again it is impossible to identify specific cause and effect. This finding is not seen when looking across all 16 village units, though it corresponds with views reported in the mid-term evaluation of Integrity Action's Sida grant:

> "When you look around in Kinango, we have less white elephant projects in our community. All the new projects are constructed with purpose and with durability in mind due to the vigilance of community monitors… There has been a massive change. Before the VOICE project, PMC were not even selected, you will wake up one morning and find a project being implemented already. Now, contractors can't start work without a community monitor." *~ Citizen monitor*

> "We have increased performance in our project delivery. Kwale County managed to become number 2 in the whole country in terms of performance." *~ Village official*

A possible impact of monitoring is also hinted at in the one-third of projects not recorded as complete and in good condition. As in the table above, of the projects that were monitored, four had been completed in poor condition and three were stalled. For the unmonitored projects, eight were completed in poor condition and only one was

stalled. This could be a sign that monitors have at least been able to prevent contractors from handing over poor-quality projects, even if they haven't been able to get all problems fixed. However, this is just one interpretation and the numbers are far too small to prove a significant link[10].

7. Are monitored projects serving a better use than others?

For the projects that were reported as completed (whether in good or poor condition), the research assistants also recorded whether the projects were in use, and if so whether their use was as originally planned. If the answer to either question is "no", brief descriptions have been provided.

Of the 250 completed projects, 221 were recorded as in use. Of these, all but one was in use as planned; the one exception was at Msulwa vocational training centre where a hostel (completed but in poor condition) had been built for the use of girls but was instead being used by boys. For the purposes of analysis, this project has been combined with the projects not yet in use as shown below. Note that there were no other completed education projects not yet in use, all other "not yet in use" projects came from the departments of health or water.

	Not yet in use as planned	In use as planned	TOTAL
Good condition	9	190	199
Poor condition	21	30	51
TOTAL	30	220	250

As we might have imagined, the percentage of good-condition projects that are in use as planned (190/199 = 95%) is much higher than for the poor-condition projects (59%). Since we know that monitoring is linked to a better condition of projects, we may therefore assume that completed projects that were monitored are more likely than unmonitored projects to be in use as planned. This can be tested as in previous sections, again using the chi-square test of independence and the table below with the ***expected values*** (for if there were no relationship), and the **observed values**:

	In use as planned	Not yet in use as planned	TOTAL
Monitored	**162** > ***158***	**18** < ***22***	180
Not monitored	**58** < ***62***	**12** > ***8***	70
TOTAL	220	30	250

In this case, $\chi^2 = 2.43$, which is less than the critical value 3.84 at $\alpha = 0.05$ and so there is no significant evidence of a relationship. The p-value is also high, at 0.49. This

[10] In total, of the nine projects in the sample that were stalled, eight were monitored and only one unmonitored. While not conclusive, this does lend weight to this theory and suggests that the case for monitoring presented elsewhere in this document could in fact be strengthened by comparing only the condition of completed projects.

means that, while the observations show monitored projects tending to be in correct use more than the unmonitored projects, we cannot say that there is a statistically-significant connection.

One of the noted limitations that may be complicating this relationship is the definition of "completed", which may in general be taken to mean that the construction has been handed over by the contractor. This does not mean they have yet been opened by the county government, and several are awaiting equipment, electricity, and/or deployment of staff. In these cases, there is still investment needed before the projects are ready to be used, but the budget spent so far has not been wasted so long as the buildings remain in good condition and the additional resources are supplied soon.

At the other end of the spectrum are projects that were once completed, but which now need major repairs or even demolition. For example, one maternity wing has been entirely burnt, and one dispensary is described as "uninhabitable" with the county having already agreed to build a new one. Both of those examples came from unmonitored projects, as did a water tank that has fallen down because "the contractor of this tank is not competent". None of the monitored projects have descriptions of such serious issues, nor do they contain any reports of deliberate vandalism – which affects three more of the unmonitored projects.

While the sample size here is low, and based on inconsistent free-text data, these give some reason to believe that projects monitored by citizens are less likely to become unusable or fall into disrepair.

8. Does it matter when monitoring begins?

Not all projects that were monitored were monitored from the beginning of implementation. The analysis aimed to explore two related questions: whether there was an advantage to monitoring being present from the beginning of project implementation, and whether monitoring that began later was still associated with better project conditions.

The data collection form asked for the years in which project implementation began and, where applicable, in which monitoring began. In 57 cases, monitoring began in a year following the start of the project – typically the next year, although sometimes there was a longer gap.

For the purposes of analysis, cases where the implementation and monitoring years match – or where monitoring began earlier[11] – have been considered as being monitored from the beginning. Of course, a year can be a long time and it is possible that there is a bigger gap between the implementation and monitoring start dates of a project where these match than in another project where these differ.

[11] This is possible as community monitors may start their involvement in the initial public participation processes before implementation begins on the project site.

For example, one project may have begun implementation in January 2019 and had monitoring begin in December 2019, while another project may have started in November 2019 and been monitored from February 2020. However, the definition used in this analysis is still of interest as a proxy indicator.

One of the monitored projects in the full sample is missing the year in which monitoring began, so the sample used here is of 258 projects as follows:

	Good condition	Poor or stalled	TOTAL
Monitored from beginning	110	20	130
Monitoring began later	47	10	57
Not monitored	41	30	71
TOTAL	198	60	258

Two final chi-square tests of independence have been conducted and the tables below present the ***expected values*** (for if there were no relationship), and the **observed values**:

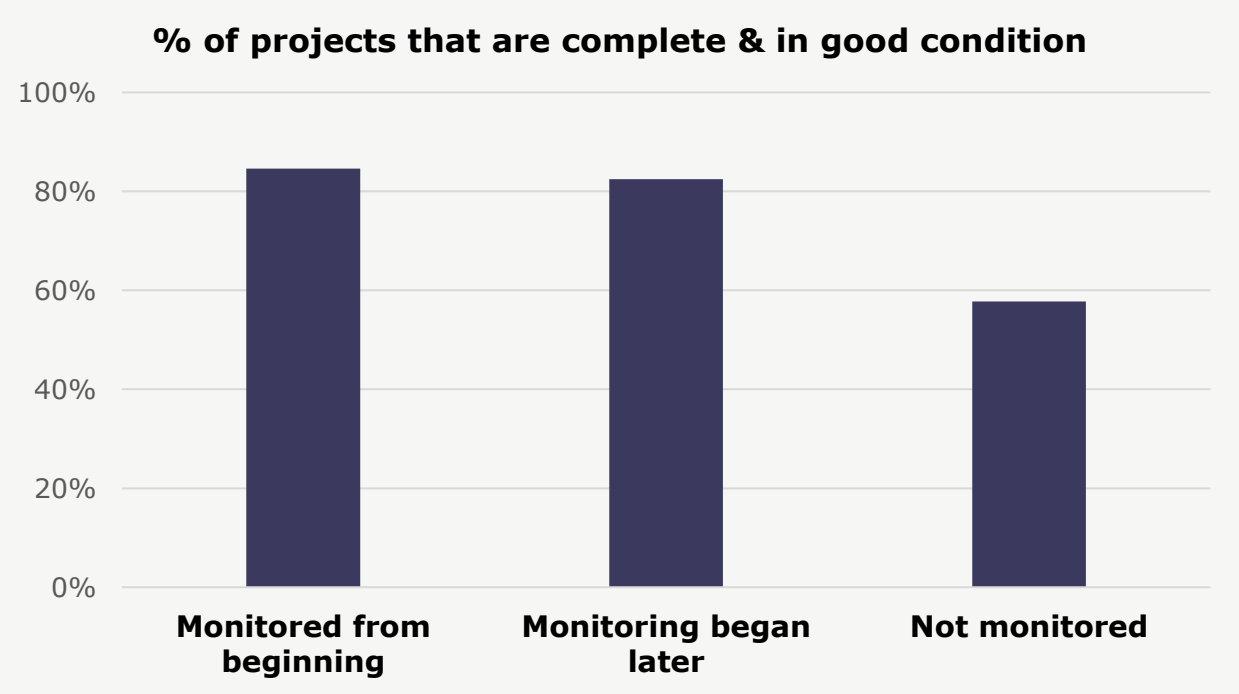
	Good condition	Poor or stalled	TOTAL
Monitored from beginning	**110** > ***109***	**20** < ***21***	130
Monitoring began later	**47** < ***48***	**10** > ***9***	57
TOTAL	157	30	187

	Good condition	Poor or stalled	TOTAL
Monitoring began later	**47** > ***39***	**10** < ***18***	57
Not monitored	**41** < ***49***	**30** > ***22***	71
TOTAL	88	40	128

The first of these tables shows a negligible difference between the expected and observed values, therefore failing to identify a significant benefit to monitoring from the beginning compared to monitoring later.

The second table provides a χ^2 value of 8.99, which is considerably greater than the critical value 3.84 at $\alpha = 0.05$, $p = 0.03$. There is therefore a **strong and statistically significant relationship between a project's condition and whether it was monitored, even in cases where monitoring did not begin in the same year as implementation.**

The differences in status are also illustrated in the following graph:



Overall, this analysis lends weight to the existing evidence that shows having community members involved in monitoring of infrastructure projects can result in better quality constructions, and furthermore that this benefit holds even if community monitors begin their activities after the start of project implementation.

Annex: Data collection form

The following form was built on KoboToolbox by KYGC and used by the assistants.

#	Question	Response options
1	Sub County	• Matuga • Kinango
2	Ward *[Response options depend on previous choice of sub county]*	• Kubo South • Tiwi • Mkongani • Waa-Ngombeni • Kinango • Kasemeni
3	Village Unit *[Response options depend on previous choice of ward]*	• Majimboni • Mangawani • Mwaluvanga • Shimba Hills • Kizibe • Mkomba • Mlafyeni • Tiribe • Mkoyo • Simkumbe • Kiteje • Ngombeni • Bofu • Minyenzeni • Gandini • Kinango
4	Department	• Education • Medical and Public Health • Water Services
5	Project Name	<free text>
6	Was project implementation monitored?	• Yes • No
7	When was the project monitored *[Only asked if Q6 = Yes]*	• 2018 • 2019 • 2020 • 2021 • 2022
8	Who monitored the project? *[Only asked if Q6 = Yes]*	• Community monitors • Project committee • Both community monitors and project committee • Other organisations /institutions
9	When did the project implementation begin?	• Before 2018 • 2018 • 2019 • 2020 • 2021 • 2022

10	Project current status	• Complete and in good condition • Complete and in poor condition • Ongoing • Stalled
11	Is the project in use? *[Only asked if Q10 = Complete]*	• Yes • No
12	If no, why? *[Only asked if Q11 = No]*	<free text>
13	Is the project being used as planned *[Only asked if Q11 = Yes]*	• Yes • No
14	If no, explain *[Only asked if Q13 = No]*	<free text>
15	Any comments?	<free text>
16	Project location	<GPS data>
17	Project photo	<photo upload>