

RESEARCH ARTICLE

Free and fair to whom? Expert and citizen assessments of election quality in twenty-two African countries

Rogers Gabula Bagiire

Department of Development Studies, Kyambogo University, Kenya

Corresponding author: Rogers Gabula Bagiire. rbagiire@kyu.ac.ug

Submitted 14 September 2025 Accepted 17 November 2025

Revised 25 April 2026 Published 20 June 2026

ABSTRACT

Purpose. Two instruments claim to measure the quality of the same election. Expert panels score it against a checklist of procedural standards; mass surveys ask the people who voted in it. Both are used, often interchangeably, as measures of electoral integrity. This article asks how far the two agree across African elections, and what follows for research that treats either as a proxy for the other.

Methods. Citizen assessments come from the Afrobarometer Round 10 country reports, which print the full response distribution for the question on the freeness and fairness of the last national election. Thirty-five countries were transcribed and each was accepted only after its two positive categories reconciled, within one percentage point, with the figure Afrobarometer published for that country in Dispatch 1128. Expert assessments come from Perceptions of Electoral Integrity 11.0. The citizen item was rescaled to 0 to 100 so that it shares the metric of the expert index, and countries were retained only where the election the survey asked about is the election the expert panel scored. Twenty-two countries survive both filters.

Findings. Citizens rate their elections 14.6 index points more favourably than experts do on average (SD 20.7; $t = 3.31$, $p = 0.003$), with 17 of 22 gaps positive. The direction is the opposite of the global pattern, where citizens are the harsher judges. More consequentially, the two measures barely track each other. Regressing the citizen measure on the expert measure gives a slope of 0.17 (SE 0.16) and an R-squared of 0.04; the slope cannot be distinguished from zero but is decisively different from one ($p = 0.0001$). Press freedom does not account for the divergence, contrary to the mechanism established on a 25-nation sample. Perceived freedom of speech enters with the sign opposite to the one preference falsification predicts.

Implications. The divergence is not well explained by citizens being misinformed or intimidated. It is better read as the two instruments measuring different things: experts score the administration of a contest, citizens appear to render a verdict on the political order that contest belongs to. Neither is thereby wrong, but they are not substitutes, and the common practice of using a survey item as a proxy for election quality, or as a validity check on expert scores, does not survive the evidence reported here.

Keywords: electoral integrity; measurement validity; Afrobarometer; Perceptions of Electoral Integrity; public opinion; African elections

1. Introduction

In October 2020 Tanzania held a general election in which the main opposition presidential candidate was detained and questioned by police four days after the vote, and in which the ruling party took 97 per cent of the directly elected parliamentary seats (Freedom House 2021). The Electoral Integrity Project's expert panel scored that contest 31 out of 100, placing it in the bottom fifth of African elections. When Afrobarometer interviewed 2,400 Tanzanians in mid-2024 and asked how free and fair that same election had been, 84 per cent said it was free and fair, and 55 per cent said it was completely so.

A researcher who needs a measure of Tanzanian election quality therefore has two numbers, 31 and 84, and no obvious way to choose. The choice is not academic. Expert indices and survey items are both used as dependent variables in work on backsliding, as independent variables in work on turnout and protest, and as validity checks on one another. If they agree, the choice is a matter of convenience. If they do not, a large body of results depends on which instrument the author happened to reach for.

The question has been asked, but not here and not at this scale. Norris (2014) compared mass perceptions with independent assessments using World Values Survey data and concluded that citizens make reasonably accurate judgements where they have access to independent media, education and party cues. Van Ham (2015) showed that expert measures themselves disagree with one another, and that they disagree most about

low-integrity and post-conflict elections. Vliegthart, van Ham, Kruikemeier and Jacobs (2024) ran the sharpest test to date, pairing a 25-nation citizen survey with matched expert items, and found citizens more negative than experts by a mean of 7.05 points, with trust in traditional media narrowing the gap and trust in social media widening it. Only one African country was in their sample. It was Kenya, and Kenya was the sole case in all 25 where citizens were the more generous judges.

The African work is thinner and narrower. Kerr (2013, 2018) examined popular evaluations of election quality in Nigeria, showing that citizens' judgements of the electoral commission's performance drive their verdict on the contest. Kerr and Lüthmann (2017) took the question cross-national but asked about trust rather than about the size of any expert-citizen divergence. Nobody, so far as I can establish, has placed African expert scores and African citizen judgements of the same elections on a common scale and asked how closely they correspond.

This article does that for twenty-two countries. The headline result is not the gap, though the gap is large and points the opposite way to the global pattern. It is the near-absence of correspondence. Across these twenty-two elections, knowing what the expert panel scored tells you almost nothing about what the electorate said, and the reverse also holds. The citizen measure moves roughly a sixth of a point for every point the expert measure moves, and the association is statistically indistinguishable from none at all.

I then test the two explanations the literature offers, and neither survives. The information account holds that citizens

in closed media environments cannot see what experts see; press freedom should therefore predict the gap, and in these data it does not. The preference falsification account holds that citizens in repressive settings report what is safe rather than what they believe; perceived freedom of speech should therefore be negatively related to the gap, and it is positively related to it. What remains is the least flattering possibility for both instruments: that they are not measuring the same object.

The article proceeds in six parts. Section 2 sets out what each instrument is built to capture and states the hypotheses. Section 3 describes the sources, the harmonisation, the reconciliation protocol and the exclusions. Section 4 reports the gap and the correspondence test. Section 5 tests the two candidate mechanisms and develops a third reading. Section 6 states what follows for research practice, and concludes.

2. What the two instruments are built to measure

2.1 The expert instrument

Perceptions of Electoral Integrity asks country specialists to score an election against eleven stages of the electoral cycle, from the legal framework and boundary delimitation through registration, campaign finance and media coverage to the count and the resolution of disputes. Version 11.0 covers 643 elections in 170 countries between July 2012 and December 2024, drawing on 5,605 expert assessments (Garnett, James and Caal-Lam 2025). The index is reported on a 0 to 100 scale.

The instrument's virtue is that it is anchored to procedure. An expert asked whether the boundaries were drawn fairly is answering a question with a defensible right answer, and the eleven stages between them cover most of the repertoire Cheeseman and Klaas (2018) identify as the practical means by which African elections are manipulated: the register, the boundaries, the location of polling stations, the state media and the courts. Its known weaknesses are equally well documented. Bollen and Paxton (2000) showed that subjective ratings of democracy carry systematic judge-specific bias related to the rater's own institutional and ideological location. Little and Meng (2024) argue that the recent global decline recorded by subjective indices is not matched by objective indicators, raising the possibility that expert panels have shifted their standards rather than the world having shifted under them. Van Ham (2015) demonstrates that the several expert datasets disagree most precisely where the stakes are highest. None of this makes the expert measure useless. It does mean that when experts and citizens disagree, it is not obvious in advance who is mistaken, and I do not assume that experts are the benchmark.

2.2 The citizen instrument

Afrobarometer asks a single question: on the whole, how would you rate the freeness and fairness of the last national election, held in a stated year, with four substantive options running from not free and fair to completely free and fair. Round 10 fielded it in 38 countries between 2024 and 2025 with 50,961 interviews (Afrobarometer 2026).

The item was never designed as a measure of electoral administration. It was designed to capture legitimacy, which is a different and in some ways more important thing. Bratton and Mattes (2001) established early that African support for democracy is more instrumental than intrinsic, responsive to what government delivers rather than to how it is constituted. Moehler (2009) showed that African citizens' evaluations of institutions separate into a critical and a submissive dimension

that do not map onto procedural quality. Weghorst and Lindberg (2013) find African vote switching driven by policy and performance rather than by procedural evaluation, which is a further reason not to expect the item to behave as an administrative audit. That is not to say African electorates are indifferent to how elections are conducted: Gutiérrez-Romero and LeBas (2020) show in a Kenyan candidate-choice experiment that a candidate rumoured to have used electoral violence loses close to nine percentage points of support. But conduct attributed to a candidate and the administration of a national contest are different objects. If the item captures something closer to consent than to compliance with standards, divergence from an expert checklist is not error; it is the instrument doing its job.

2.3 Why they might diverge, and three hypotheses

Three accounts are available.

The **information account** is the one Vliegthart and colleagues (2024) test and support. Citizens can only judge what they can observe, and what they can observe is mediated. Where the press is constrained, irregularities documented by experts never reach the electorate, so citizen ratings stay high while expert ratings fall. Kerr and Lüthmann (2017) develop the same logic for trust in African elections.

The **fear account** holds that the divergence is not perceptual but expressive. Where speaking freely is risky, respondents give the answer that is safe to give. Kuran (1995) supplies the general theory; Robinson and Tannenber (2019) demonstrate the effect experimentally in a closed regime; and Friesen (2022), working directly in Afrobarometer data across 34 African countries, estimates that ruling-party support is overstated by more than ten percentage points in Sudan and Uganda and by rather less elsewhere, with the strongest markers being perceived government sponsorship of the survey and residence in a community dominated by the ruling party.

The **construct account** holds that the two instruments ask about different objects that happen to share a name. The expert is scoring an administrative process. The citizen, asked whether an election was free and fair, may be answering a broader question about whether the political order permits people like them to have a say. On this reading the divergence carries no implication that either side is mistaken.

These yield testable predictions.

Hypothesis 1. Across African elections, citizen assessments exceed expert assessments of the same election.

Hypothesis 2. The citizen measure and the expert measure correspond: the slope of one on the other is positive and appreciably different from zero.

Hypothesis 3a (information). The gap is larger where press freedom is lower.

Hypothesis 3b (fear). The gap is larger where citizens report less freedom to say what they think.

Hypotheses 3a and 3b make opposite empirical demands on the same data, which is what makes them worth testing together.

3. Data and measurement

3.1 Sources

The citizen data are the Afrobarometer Round 10 country Summary of Results reports, one per country, which print the national frequency distribution for every question. These were used rather than the merged microdata because they carry the full distribution, they name the election the question refers to, and they are the form in which Afrobarometer itself publishes country results. The expert data are the PEI 11.0 index scores and election dates as published in Table 1 of the Electoral Integrity Global Report 2025 (Garnett, James and Caal-Lam 2025). Press freedom is the Reporters Without Borders World Press Freedom Index 2026 country score, on a 0 to 100 scale (RSF 2026). Freedom of speech and trust in the electoral commission are taken from the same Afrobarometer country reports as the dependent variable, which means the mechanism measures come from the same respondents in the same interviews.

3.2 A reconciliation protocol

Transcribing figures from published country reports invites error, so every country was subjected to an external check before use. Afrobarometer's Dispatch 1128 reports, for each of 38 countries, the share of respondents rating the last election free and fair with minor problems or completely free and fair (Afrobarometer 2026). For each country I added the corresponding two categories from that country's own Summary of Results and compared. A country entered the dataset only if the two agreed within one percentage point. Kenya illustrates: the country report gives 31.5 per cent and 24.7 per cent, summing to 56.2, against a published 56.

Thirty-five countries were transcribed and all thirty-five reconciled. One, Cabo Verde, reconciled only after the four response categories were read in the reverse of the order first extracted; that country is flagged in the results and the analysis is reported with and without it. Three further countries, Benin, Gabon and Guinea-Bissau, have no Round 10 Summary of Results published and could not be attempted.

3.3 Harmonisation

The citizen item is a four-point scale and the expert index runs 0 to 100. I assign the four substantive categories the values 0, 33.3, 66.7 and 100, and take the mean over valid responses, excluding don't know and refusals. This places the citizen measure on the expert measure's metric and preserves the information in the middle categories that a simple top-two percentage discards. The gap is the citizen mean minus the expert score, in index points. Freedom of speech and electoral-commission trust are rescaled the same way.

3.4 The matching problem, and what it costs

The two instruments must refer to the same election or the comparison is empty. This turns out to be a more serious constraint than it appears. The Afrobarometer item names the election it asks about; PEI Table 1 reports the most recent election it has scored for each country. These coincide less often than one might assume, because survey rounds and electoral calendars are not synchronised.

Of the 35 reconciled countries, 22 match. Thirteen do not, and the mismatches are substantial rather than marginal: Botswana's respondents were asked about 2019 while the panel scored 2024; Ghana's about 2020 against a scored 2024; Liberia's about 2017 against 2023; Mauritania's about 2024

against a scored 2018. Comoros and Seychelles fail for the opposite reason, the survey naming an election more recent than any the panel has scored. Zambia is a case of its own: its Round 10 report prints the question as referring to an election "held in 2011", which is almost certainly an uncorrected template error for 2021, and I exclude it from the primary analysis on that ground while reporting it in the robustness table.

I record this at length because merging expert election scores onto survey data is done casually in this literature, and thirteen of thirty-five is a high failure rate for a step that is usually not reported at all.

4. How far do the two measures agree?

4.1 The gap

Table 1 reports the twenty-two matched countries.

Table 1. Citizen and expert assessments of the same election

Country	Election	Citizen	Expert	Gap
Eswatini	2023	74.1	23	+51.1
Tanzania	2020	81.1	31	+50.1
Madagascar	2024	74.9	36	+38.9
Chad	2024	53.4	18	+35.4
Zimbabwe	2023	53.0	18	+35.0
Senegal	2024	88.9	54	+34.9
Sierra Leone	2023	66.9	36	+30.9
Uganda	2021	54.9	31	+23.9
Mali	2020	58.6	39	+19.6
Guinea	2020	46.5	29	+17.5
Cote d'Ivoire	2021	62.4	48	+14.4
Congo Republic	2022	36.2	26	+10.2
Mozambique	2024	36.5	29	+7.5
Cameroon	2020	41.5	36	+5.5
Lesotho	2022	76.2	72	+4.2
Nigeria	2023	44.5	43	+1.5
Angola	2022	40.8	40	+0.8
Cabo Verde	2021	67.1	77	-9.9
Morocco	2021	52.5	63	-10.5
Kenya	2022	52.4	63	-10.6
Malawi	2020	45.3	59	-13.7
South Africa	2024	52.9	68	-15.1

Notes: both measures run 0 to 100. The citizen figure is the mean of the four-point Afrobarometer item over valid responses; the expert figure is the PEI 11.0 index for the same election. Cabo Verde's response distribution required re-ordering to reconcile with the published national figure and is treated as provisional.

The mean gap is **+14.62 points** (SD 20.71, median +12.28), and it is distinguishable from zero ($t = 3.31$, $p = 0.003$). Seventeen of twenty-two countries show citizens rating the election above the expert panel. The five exceptions are Cabo Verde, Morocco, Kenya, Malawi and South Africa, and their position in the distribution is the first thing worth noticing: they are, with Morocco the arguable exception, among the better-performing electoral democracies in the set. Citizens are the harsher judges precisely where experts award the highest marks.

This reverses the global pattern. Vliegenthart and colleagues (2024) found citizens more negative than experts by a mean of 7.05 points across 25 nations, with France at -11.51 the most negative and Kenya at +0.94 the only positive case in the sample. Their instrument and mine are not the same, so the two numbers should not be differenced. But the sign of the

African average is unambiguous, and their Kenyan anomaly looks less anomalous once the continent is in view.

4.2 The correspondence test, which matters more

The gap correlates strongly and negatively with the expert score ($r = -0.710$, $p = 0.0002$). That association is partly built in, since the expert score appears on both sides, and I do not lean on it. The substantive question is whether the two measures track one another at all, which is answered by regressing the citizen measure on the expert measure. Under perfect correspondence the slope is one.

Figure 1 plots the twenty-two elections against the line of perfect correspondence.

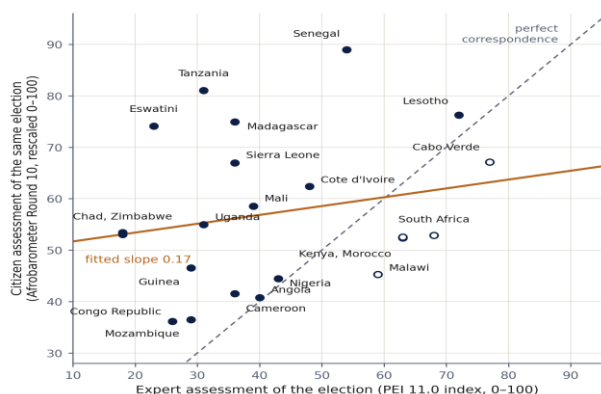


Figure 1. Citizen against expert assessments of the same election, twenty-two African countries. Filled markers are countries where citizens rate the election above the expert panel, hollow markers those where they rate it below. The dashed line is perfect correspondence; the solid line is the fitted relationship. Chad and Zimbabwe coincide almost exactly, as do Kenya and Morocco; each pair carries one label.

The fitted slope is **0.172** (SE 0.162), with an intercept of 49.97 and an R^2 of 0.042. The slope is not distinguishable from zero ($p = 0.29$) and is decisively distinguishable from one ($t = -5.10$, $p = 0.0001$). The raw correlation between the two measures is $+0.204$ and does not reach conventional significance ($p = 0.36$). Hypothesis 2 fails.

The failure is not an artefact of restricted range. The citizen measure has a standard deviation of 14.9 points against the expert measure's 17.7, so both instruments discriminate; they simply do not discriminate along the same dimension. Nor does it turn on any single case. Excluding the provisional Cabo Verde reading gives a slope of 0.146 and a mean gap of $+15.79$. Adding Zambia back on a corrected 2021 election year gives 0.163 and $+15.69$. Dropping the election-matching requirement altogether and using all 34 reconciled countries with a PEI score gives 0.249 and $+11.91$. Every specification rejects a slope of one at $p < 0.001$, and none can reject a slope of zero.

Hypothesis 1 is supported and Hypothesis 2 is rejected. The second result is the more consequential, because a systematic gap of known size could be corrected for, whereas a near-zero correspondence cannot.

5. Why do they diverge?

5.1 The information account fails

Press freedom does not explain the divergence. The bivariate association between the RSF score and the gap is negative, as the account predicts, but modest and short of conventional significance ($r = -0.429$, $p = 0.053$; Spearman $\rho = -0.323$, $p = 0.15$). Once the expert score is controlled, the RSF coefficient collapses to $+0.011$ (SE 0.397, $p = 0.98$) and

remains insignificant in the fuller specification ($+0.362$, SE 0.355, $p = 0.31$). Table 2 reports the models.

Table 2. The gap regressed on the expert score and the two candidate mechanisms

	Model 1	Model 2	Model 3
Expert score	-0.828 (0.162)	-0.831 (0.225)	-1.175 (0.220)
Press freedom (RSF)		+0.011 (0.397)	+0.362 (0.355)
Freedom of speech			+0.719 (0.268)
Intercept	+49.97 (7.93)	+49.53 (16.95)	+1.62 (23.73)
n	22	22	22
Adjusted R^2	0.479	0.451	0.590

Notes: OLS with heteroskedasticity-robust (HC3) standard errors in parentheses. ** $p < 0.01$, *** $p < 0.001$. The dependent variable is the citizen assessment minus the expert assessment, in index points.

This is a direct non-replication of the mechanism Vliegthart and colleagues (2024) establish on a mostly OECD and Latin American sample. I would not want to overstate it: their design measures individual-level media trust, mine measures a country-level press-freedom score, and a null at the country level does not refute an individual-level finding. What can be said is that the aggregate version of the mechanism, which is the version most comparative work would use, does no work in these data.

5.2 The fear account fails, and fails informatively

Perceived freedom of speech is unrelated to the gap bivariate ($r = -0.005$, $p = 0.98$). Conditional on the expert score and press freedom it is significant, but with the wrong sign: $+0.719$ (SE 0.268, $p = 0.007$). Among countries whose elections the experts scored equally, those whose citizens report feeling freer to speak give the *more* generous verdict on the election, not the less generous one. Preference falsification predicts the reverse.

Tanzania makes the point concretely. Its 2020 election scores 31 with the experts. Its citizens rate it 81. And 88 per cent of Tanzanian respondents say they feel somewhat or completely free to say what they think, the highest figure among the twenty-two. If Tanzanians were giving the safe answer about their election, one would expect them to be giving the safe answer about their freedom to speak as well, which would depress rather than inflate that figure. Compare Comoros, where 75 per cent say they are not very or not at all free to speak, and where citizens rate the 2025 election at the bottom of the whole distribution.

Two qualifications are owed. First, self-reported freedom of speech is itself vulnerable to the bias it is being used to detect, so this test is weaker than it looks; a list experiment or the machine-learning correction Friesen (2022) develops would be a stronger instrument, and neither is available at country level for Round 10. Second, Friesen's own estimates are drawn from Round 6, a decade before the elections examined here, which is why I have not used them as a covariate. What the present evidence supports is the narrower claim that the gap does not behave the way a fear account requires, not the broader claim that self-censorship is absent from African survey data.

5.3 What is left

If citizens are neither misinformed nor silenced, the remaining possibility is that they are answering a different question from the one the experts answer.

The pattern in Table 1 is consistent with this. The five countries where citizens are harsher than experts are countries with functioning oppositions, alternation in office and a record of competitive elections. Their citizens have seen elections work and have a standard against which to find a particular contest wanting. South Africa's 2024 election, which experts scored 68 and which ended the African National Congress's parliamentary majority, was rated 52.9 by South Africans. Kenya's 2022 contest, scored 63, was rated 52.4. The Kenyan case is instructive because the standard is documented: Ferree, Gibson and Long (2014) show that Kenyan voters' beliefs about irregularities in 2013 were structured and consequential rather than diffuse, and Kerr and Wahman (2021) show that African publics update on judicial rulings about elections. An electorate whose Supreme Court annulled a presidential election in 2017 has a demanding and specific benchmark, and the expert panel's 63 does not meet it. This is the critical-citizen dimension Moehler (2009) identified, and it is what one would expect of an electorate judging an election against experience rather than against a checklist.

At the other end, Eswatini's 2023 election returned a parliament to an absolute monarchy in which political parties cannot contest, and the expert panel scored it 23. Eswatini's citizens rated it 74.1. It is difficult to read that as a claim about ballot administration. It is easier to read as an answer to a question about whether the process was acceptable, asked of people for whom the alternative to that process is not a better election but no process at all. Levitsky and Way (2010) make the related point that elections in hybrid and closed regimes perform functions other than selecting a government, and a citizen evaluating one of them may reasonably be evaluating what it is for rather than how well it was run.

The construct account does not require that either instrument is defective. It requires only that "free and fair" means something different to a country specialist scoring eleven stages of an electoral cycle and to a citizen who voted in it. Adcock and Collier (2001) would call this a failure of measurement validity rather than of reliability: both instruments may measure their own construct well while measuring each other's badly.

I cannot adjudicate this against the alternatives with these data, and I do not claim to. Country-level aggregates cannot distinguish a compositional story from an individual-level one, and the construct account is the residual, supported by the failure of its rivals and by the shape of the pattern rather than by a direct test. A design that could adjudicate is set out below.

6. What follows, and conclusion

Three practices in the existing literature do not survive these results.

Substituting one measure for the other. A slope of 0.17 with an R^2 of 0.04 means that a study using the Afrobarometer item as a proxy for election quality, or an expert index as a proxy for what citizens believe, is not approximating the other quantity. It is measuring something else and calling it by the other's name. This bears directly on work that uses survey measures of democracy as outcome variables in studies of backsliding, which is now a substantial literature. The stakes are not trivial. The V-Dem Institute records forty-four countries autocratizing in 2025, twelve of them in sub-Saharan Africa (Nord et al. 2026), and the incremental, legalistic character of contemporary erosion that Bermeo (2016) describes is precisely the form least likely to register in a single summary survey item. A study that tests for

backsliding using such an item, in the region where most autocratization is occurring, may be measuring consent while believing it is measuring procedure.

Treating citizen ratings as a validity check on expert scores, or the reverse. Convergent validation requires that the two instruments be attempting the same target. If the construct account is right, agreement between them would be evidence of nothing in particular and disagreement is evidence of nothing either.

Merging expert election scores onto survey data without reporting the match. Thirteen of thirty-five countries here fail the requirement that the two sources refer to the same election, and the failures are of five and seven years rather than of months. Any study performing this merge should report the matching rate and the exclusions.

Four limitations bound what has been shown. The analysis is at country level with twenty-two observations, which is enough to reject a slope of one decisively but not enough to model the sources of heterogeneity. The mechanism measures are country aggregates and cannot rule out individual-level versions of either account. One country's response distribution, Cabo Verde's, was reconstructed to reconcile with the published national figure and should be treated as provisional, though the results hold without it. And the expert measure is itself contested: Bollen and Paxton (2000) and Little and Meng (2024) give reasons to doubt that expert panels are the fixed point against which citizens should be assessed, and nothing here establishes that they are.

Two designs would take this further. The first is an individual-level replication using Afrobarometer microdata with country fixed effects, asking whether the respondents whose assessments diverge most from the expert score are the better educated and the more media-exposed, as the information account requires, or the more regime-aligned, as a partisan-rationalisation account would predict. The second, and the more decisive, is a split-ballot experiment: ask half of a national sample the standard freeness-and-fairness question and the other half a battery of specific procedural questions of the kind the expert panel answers, on the register, the count, the media and the resolution of disputes. If the construct account is right, the procedural battery will track the expert score and the summary item will not, within the same population and the same week.

Citizens in twenty-two African countries rated their most recent national elections, on average, 14.6 points more generously than the specialists who scored the same contests, and what they said carried almost no information about what the specialists recorded. The temptation is to treat that as a finding about African publics being poorly informed or unfree to speak, and the evidence assembled here does not support either reading. The more uncomfortable possibility is that a question the discipline has been asking for twenty-five years, and reading as a measure of electoral integrity, has been answering a different question all along.

References

- Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, 95(3), 529–546.
- Afrobarometer. (2026). *Africans want elections but don't trust electoral management bodies to ensure they're free and fair*. Dispatch No. 1128. Accra: Afrobarometer.

- Bermeo, N. (2016). On democratic backsliding. *Journal of Democracy*, 27(1), 5–19.
- Bollen, K. A., & Paxton, P. (2000). Subjective measures of liberal democracy. *Comparative Political Studies*, 33(1), 58–86.
- Bratton, M., & Mattes, R. (2001). Support for democracy in Africa: Intrinsic or instrumental? *British Journal of Political Science*, 31(3), 447–474.
- Cheeseman, N., & Klaas, B. (2018). *How to rig an election*. New Haven: Yale University Press.
- Ferree, K. E., Gibson, C. C., & Long, J. D. (2014). Voting behavior and electoral irregularities in Kenya's 2013 election. *Journal of Eastern African Studies*, 8(1), 153–172.
- Freedom House. (2021). *Tanzania: Freedom in the world 2021*. Washington, DC: Freedom House.
- Friesen, P. (2022). *Measuring the impact of self-censorship on political party support in Afrobarometer data using machine learning*. Afrobarometer Working Paper No. 192. Accra: Afrobarometer.
- Garnett, H. A., James, T. S., & Caal-Lam, S. (2025). *Electoral integrity global report 2025*. The Electoral Integrity Project.
- Gutiérrez-Romero, R., & LeBas, A. (2020). Does electoral violence affect vote choice and willingness to vote? Conjoint analysis of a vignette experiment. *Journal of Peace Research*, 57(1), 77–92.
- Kerr, N. (2013). Popular evaluations of election quality in Africa: Evidence from Nigeria. *Electoral Studies*, 32(4), 819–837.
- Kerr, N. (2018). Election-day experiences and evaluations of electoral integrity in unconsolidated democracies: Evidence from Nigeria. *Political Studies*, 66(3), 667–686.
- Kerr, N., & Lührmann, A. (2017). Public trust in manipulated elections: The role of election administration and media freedom. *Electoral Studies*, 50, 50–67.
- Kerr, N., & Wahman, M. (2021). Electoral rulings and public trust in African courts and elections. *Comparative Politics*, 53(2), 257–290.
- Kuran, T. (1995). *Private truths, public lies: The social consequences of preference falsification*. Cambridge, MA: Harvard University Press.
- Levitsky, S., & Way, L. A. (2010). *Competitive authoritarianism: Hybrid regimes after the Cold War*. Cambridge: Cambridge University Press.
- Little, A. T., & Meng, A. (2024). Measuring democratic backsliding. *PS: Political Science and Politics*, 57(2), 149–161.
- Moehler, D. C. (2009). Critical citizens and submissive subjects: Election losers and winners in Africa. *British Journal of Political Science*, 39(2), 345–366.
- Nord, M., Altman, D., Fernandes, T., Good God, A., & Lindberg, S. I. (2026). *Democracy report 2026: Unraveling the democratic era?* (10th ed.). Gothenburg: V-Dem Institute, University of Gothenburg.
- Norris, P. (2014). Public perceptions of electoral integrity. In *Why electoral integrity matters* (pp. 91–110). Cambridge: Cambridge University Press.
- Reporters Without Borders. (2026). *World press freedom index 2026*. Paris: RSF.
- Robinson, D., & Tannenbergh, M. (2019). Self-censorship of regime support in authoritarian states: Evidence from list experiments in China. *Research and Politics*, 6(3), 1–9.
- van Ham, C. (2015). Getting elections right? Measuring electoral integrity. *Democratization*, 22(4), 714–737.
- Vliegthart, R., van Ham, C., Kruikemeier, S., & Jacobs, K. (2024). A matter of misunderstanding? Explaining (mis)perceptions of electoral integrity across 25 different nations. *Public Opinion Quarterly*, 88(SI), 495–515.
- Weghorst, K. R., & Lindberg, S. I. (2013). What drives the swing voter in Africa? *American Journal of Political Science*, 57(3), 717–734.

Declarations

Funding. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interests. The author declares none.

Data availability. The coded country dataset, the reconciliation checks against Dispatch 1128 and the analysis script that produces every estimate reported in Sections 4 and 5 are available from the author on request and will be deposited in a public repository on acceptance.