

January 2026

Reality Report

A transparent look at how we protect data quality and deliver real accuracy. From panel to methodology.

2026

Foreword

By Stephan Shakespeare

I co-founded YouGov in 2000 to do something simple but transformative: **use the internet to measure opinion continuously and at scale.** Instead of occasional, expensive snapshots, we wanted a connected, living dataset—updated daily, comparable over time, and precise enough to guide real-world decisions. That meant building our own panel, using technology to help us better understand what people think.

As we've strived to create a virtuous circle of trust, the industry has suffered a vicious one. All of us have seen how for many, over the years, a poor participant experience has driven good panelists away, magnified the effect fraud and led to a decline in quality. Panelists are now brokered in real time—their provenance concealed, their veracity doubted—fueling skepticism and cynicism in the industry and beyond.

From the outset, we treated our panel as our core asset. Everything else followed from that. A strong panel is not a list of email addresses; it is a community with known provenance, where people feel their time is respected and their voice has weight. They see the value of being part of something larger, and the impact of their opinions and actions on elections, policies, and the decisions brands make every day.

That dynamic strengthens itself: a good experience keeps people engaged, which leads to better data for clients, which in turn builds trust in the results—and makes being part of YouGov something people actively value.

We chose a different path, and it has proved to be the right one.

Treating our panel as an asset—owning the relationship and every aspect of the process from recruitment and verification to sampling, fieldwork and validation—means we can take the

right decisions and act with care. We invest in a good user experience and never rent them out to other providers. We do our utmost to deter and detect bad actors. And we develop systems and methods that ensure we deliver data our clients can trust.

The results show up where it counts. When elections, crises or major commercial decisions demand a firm read of public sentiment, our numbers are used because people know they reflect the reality of the world around them. That is the result of thousands of operational choices we make in favor of quality.

This collection (this e-book, these pages, these articles) sets out these choices. It details how we think about quality, how we design the participant journey, how we verify provenance, how we sample and weight, and how we test and correct. These are not standalone techniques or disjointed processes. They are the inseparable parts of an integrated system backed by a long-term, coherent line of thinking.

Our approach to panel also fuels our innovation. Leaps forward such as our groundbreaking daily brand tracking product BrandIndex in 2005 and our market-moving MRP modelling in the 2017 UK general election were only possible because of the relationship we have with our panel.

Now, as AI opens up new possibilities, it will prove pivotal again and we will continue to use new technology to help us better understand what people think.

We're publishing this report for two reasons. First, we're proud of it and of what it represents in terms of our accountability in showing the reality of the world around us. Our approach is nuanced, data-driven, and under our control. We put quality first. We should show our workings.

Second, we see it as a call to arms for the industry to practice what it preaches. If transparency is the logical outcome of confidence in one's approach, we encourage others to do the same.

Our approach will continue to evolve—it has to. We are not perfect, and our industry and the world around us are changing rapidly. As we make improvements, and even when we make mistakes, we will share them with you. That is how trust in data is earned—and how it is kept.



Stephan Shakespeare
Chief Executive Officer

Pillars of quality

“Can I trust the data?”

That is the question every client asks, explicitly or implicitly, when they commission research. It's the question that sits at the heart of our work. Clients rely on findings they can't afford to get wrong. They need to know it reflects reality. Clients may not frame it in methodological or operational terms, but “Can I trust the data?” really asks three things:

Are you talking to the right people?

Are those people real?

Are their answers correct?

At YouGov, we organize those concerns into three elements: representativeness, integrity, and accuracy. Each addresses a different concept; together they provide a framework for deciding whether the data and insight we provide are worthy of trust. Together, they are our Pillars of Quality.

Representation

Representativeness is the first, most fundamental requirement of survey research. If the people you interview don't reflect the population you're trying to understand, the results will be misleading—no matter how engaged the panelist.

Different studies have different needs and tolerances: a daily brand tracker, an election poll, and a niche usage and awareness study each define “the right people” differently. Who should be invited? In what proportions? Because we own our panel and recruit directly across multiple channels, we can maintain balanced participation across demographic, geographic, and behavioral profiles. Owning that relationship also gives us the explicit permission to collect, link, and use multiple kinds of data—so we can know exactly who we are speaking to.

There are four places in which we operate to maintain the representativeness of our research samples.

1. Recruitment – building the right universe

Representativeness starts with the people in our panel. We recruit widely and deliberately, validating identities and locations up front, so our starting population is already a strong match to the audiences our clients need. The data we capture at registration create the foundation of our “identity spine,” a persistent record (in some cases spanning 25 years) we use throughout the panelist lifecycle to manage sampling, data, and authenticity.

› **Global coverage and local detail.** Our proprietary panels cover more than 60 markets, with local language splits where relevant. This ensures we can set sampling frames that reflect real-world populations.

› **Multiple recruitment channels.** Panelists are recruited through a range of targeted channels, including search, social, and affiliates. Sign-up is cross-platform, on our app and via our website.

› **Verification and fraud prevention from day one.** Every new joiner is assigned an initial risk score and, at minimum, subject to email validation, double opt-in, device fingerprinting, and multi-source geolocation/VPN checks before they can access a client survey. Any account identified as high risk for geographical or duplicate fraud must complete ID verification using a unique government-issued document through our trusted specialist technology partner. (See the Integrity section for more information on how we manage fraud detection over time.)

› **Balanced composition.** We monitor panel demographics against census and official benchmarks and run targeted recruitment to fill gaps in under-represented groups.

› **Data transparency.** We publish “right-to-contact” panel counts twice yearly, separate from feasibility estimates, so clients have a clear view of our starting universe.





2. Profiling – understanding who we have

The richness of our profile data underpins both precise targeting in fieldwork and accurate adjustment after data collection. We maintain multiple streams of information for every panelist, and because these are our panelists—not anonymous or brokered traffic—we can connect these streams with confidence. That richness not only improves sampling precision but also makes analysis more powerful and reduces survey burden by not asking what we already know.

- › **Core demographics.** Variables such as age, gender, region, education, income, social grade, and ethnicity (where applicable) are collected for all panelists.
- › **Declared attributes.** Profiling surveys capture interests, attitudes, behaviors, media use, and category consumption, refreshed at regular intervals (typically every 3-6 months) to keep them current.
- › **Observed and paradata.** Device information, brand, model, OS, browser, connection type, languages, breakoff patterns, and geolocation provide behavioral context and help validate declared information.
- › **Automated refresh.** Core demographics are updated automatically from survey data, typically every 3–6 months, reducing the need to re-ask and lowering survey burden.
- › **The identity spine.** Every panelist's record begins at registration with verified identifiers and technical markers. Over time, we link new data points—updated demographics, declared attributes, behavioral and paradata signals—to this persistent record. This allows us to validate respondents at multiple points, spot anomalies

in profile or behavioral data, and maintain a consistent, trustworthy identity across years of participation.

- › **Governance and privacy.** All data handling follows our global privacy framework, applying relevant data protection practices in each market and ensuring consent and handling as required by law.

3. Data collection – sampling and fieldwork

Owning the panel means we control who is invited, when, and under what quota rules—avoiding the distortions that come from open, self-selecting access.

- › **Active sampling.** Panelists are invited to take a survey and then get given one based on what's in field and what they qualify for at the time sampling is run. When they click the invite, we check fieldwork again and they are allocated to the survey that needs them most at that point in time.
- › **Quota-driven allocation.** Proprietary internal routing assigns respondents to surveys based on project quotas, not personal choice. We can manually prioritize urgent projects, but allocation is predominantly automated against the sampling frame.
- › **Screening from profile, not survey.** Where possible, eligibility is determined from existing profile data rather than lengthy in-survey screeners.
- › **Live field monitoring.** We track fill rates, dropout points, and anomalous speeds in real time, with rule-based mid-field adjustments to keep samples aligned.

- › **Quality safeguards in-survey.** Attention checks, soft launches, and consistency tests are standard; suspect cases are routed to our Response Quality Survey for deeper assessment.
- › **Participant experience.** Every project collects post-survey satisfaction feedback, helping us maintain engagement and minimize bias from fatigue.

4. Weighting – aligning to the target population

Even well-designed samples benefit from statistical alignment to population benchmarks. Our weighting process ensures the achieved sample matches the characteristics of the population of interest.

- › **Authoritative sources.** We weight to census data, large-scale probability surveys, election results, and official statistical estimates.
- › **Market-appropriate targets.** Core demographic variables are always included; political work adds past vote and political attention, and some geographies require additional dimensions such as ethnicity or country of birth.
- › **Method choice.** We use raking and calibration methods and will apply more sophisticated model-based approaches such as multilevel regression with post-stratification (MRP) for small-area or complex estimations (like with polling).

- › **Weight stability.** Caps and trimming rules keep variance under control; in long-running trackers, we monitor for stability over time and rebase when needed.
- › **Proven outcomes > better methods.** Our methods are tested in the most demanding contexts—daily brand tracking and public polling—where we have a track record of stability and predictive accuracy, such as correctly projecting 92% of seats in the 2024 UK general election.



The richness of our profile data underpins both precise targeting in fieldwork and accurate adjustment after data collection.



Integrity

Even a sample that looks representative on paper can be undermined if participants aren't genuine. Integrity means ensuring that respondents are real, unique, and the right fit for the research being done.



Concerns about integrity have exploded in recent years. The rise of organized survey farms, geo-masking through VPN and proxy servers, duplicate accounts, bots, and now AI-assisted responses makes detection a vital, continuous discipline. Our approach is layered across the panelist lifecycle—from registration to participation to redemption—combining identity checks, device fingerprinting, multi-source geolocation, real-time threat scoring, and payout oversight to ensure bad actors do not slip through the net. We use a wide array of signals that include device fingerprinting and multi-source geolocation checks, speed checks, in-survey attention/consistency checks, and open-end quality modeling to determine if people are who they claim to be.

But rigor alone isn't enough. Fraud prevention at scale requires nuance. Our direct relationship with panelists allows us to make decisions with precision and accountability—to protect the data without eroding the trust that allows us to collect it.

There are four places we operate to maintain integrity.

1. Registration & onboarding

The most efficient fraud control is prevention. By stopping most bad actors before they ever reach a client survey, we make fieldwork more predictable, improve the reliability of our feasibility estimates, and reduce after-the-fact cleaning.

Our onboarding doesn't just "let people in"; it establishes a persistent identity spine that links every questionnaire response with every datapoint we can observe about their device, location, and behavior. That spine is the anchor by which we decide who to invite, who to verify, and—if necessary—who to remove. The experience is intentionally asymmetric: low-risk





people will have a smoother experience, while riskier cases meet progressively harder checks that quickly surface bad behavior without punishing genuine newcomers.

Operational controls:

- › **Identity & access:** email validation, double opt-in checks, email domain screening, contact detail de-duplication (name, phone, address, email), and double-keying on registration/login.
- › **Device & location intelligence:** banking-grade device fingerprinting; multi-source geolocation with VPN/proxy/residential proxy detection; cookie/machine ID checks and cookie blocklist matching.
- › **Real-time threat scoring:** partner APIs assess risk multiple times at sign-up; high-risk joiners face added friction, including additional verification questions or routing to the Fraud Detection/Response Quality Survey.
- › **Structured onboarding & RQS:** early cooperation/open-end quality checks establish a baseline; RQS uses hundreds of technical, behavioral, and open-end indicators to assess integrity from the outset for suspicious accounts.
- › **Escalation when needed:** secure identity verification (passports, national identity cards, driver's licenses) for geo/ duplicate-risk accounts; bot and AI detection on open-end responses where appropriate.

fieldwork on track. Suspect cases are routed to deeper assessment rather than bluntly removed, preserving representation while engaged respondents proceed without interruption.

Operational controls:

- › **Technical re-checks:** device fingerprints and geolocation captured on each survey; VPN/proxy/residential proxy use flagged; cookie/machine ID consistency checked against account history.
- › **Behavioral signals:** calibrated speeding thresholds, skip/straight-line detection, copy/paste pattern detection, plus soft-launch reviews.
- › **Attention & cooperation checks:** attention-check questions embedded in-survey; non-passers may be routed to RQS for full quality scoring rather than immediate removal.
- › **Consistency & plausibility:** cross-item checks and cross-survey history checks (e.g., OS/device mismatches, illogical or contradictory combinations, improbable location changes).
- › **Open-end quality at scale:** model-based scoring for responsiveness/relevance; LLM detection for AI-generated text; bot detection for nonsensical or duplicate content.
- › **Proprietary methods:** Some in-survey methods which we normally associate with improving the accuracy of our data—such as our Awareness Cross-Entropy (ACE) for brand tracking or our in-house questionnaire design safeguards—can also act as integrity signals when the anomalies they detect are severe enough to indicate fraud.

2. In-survey safeguards

Every live project is a fresh opportunity to confirm that the person in the survey is still the right person, on the right device, in the right place. A combination of adaptive real-time technical and behavioral checks block contamination before quotas fill and keep

3. Lifecycle monitoring & scoring

A single survey rarely tells the whole story. Connecting signals across projects and time is an essential practice to ensure the panel stays healthy and predictable without sacrificing scale. This evidence-based posture supports recontact and longitudinal work, steadies feasibility, and protects the quality of our normed (tracking) and sensitive (polling) studies. Our approach blends rule-based guardrails with machine-learning models that detect evolving threats so we can distinguish an occasional slip from a true pattern of risk.

Operational controls:

- › **Account-level context:** tenure, recruitment source, prior verifications, historical flags, historical redemption patterns, and time-weighted risk scoring (recent activity weighted more heavily).
- › **Response Quality Score (RQS):** hundreds of indicators (technical, behavioral, open-end, cooperation) roll into a respondent-level score; administered to both randomly selected panelists and those flagged in other checks.
- › **Rules and models:** fixed thresholds where appropriate; adaptive cutoffs by sector/market where risk is uneven; ML classifiers trained on labeled high- and low-quality cases.
- › **Targeted actions:** invite, quarantine, or remove—chosen from evidence, not guesswork. “Quarantine” status prevents survey access while preserving the account for investigation.
- › **Feedback loops:** results from lifecycle monitoring feed upstream into onboarding risk models, in-survey dynamic checks, and payout risk scoring.

- › **Cross-account detection:** device, IP, cookie/machine ID, and payout-detail re-use detection across accounts.



A single survey rarely tells the whole story. Connecting signals across projects and time is an essential practice to ensure the panel stays healthy and predictable without sacrificing scale.



4. Redemption & payout controls

Fraud usually has a financial motive. Treating redemption as the last integrity checkpoint blocks bad actors before cash-out, depresses the “ROI” of fraud, deters repeat attempts, and ensures incentives go to genuine panelists. Crucially, payout-stage intelligence feeds back into earlier stages, helping surface new rings and tactics sooner. If a fraudster managed to behave plausibly long enough to reach a reward, this is where financial signals help us shut the door.

Operational controls:

- › **Account reviews:** account age/history, recruitment source, earnings patterns, redemption frequency and cadence, payout velocity monitoring.
- › **Financial controls:** bank-detail validity and geo checks; bank account reuse detection across accounts; physical goods redemption review (if applicable).
- › **Ecosystem signals:** gift-card fingerprinting via fintech partners, including cross-market device and redemption matching; sanctions/AML screening via specialist technology partners.
- › **Hold/review logic:** suspect redemptions paused for verification; confirmed fraud de-monetized and removed; linked accounts investigated and, if necessary, actioned across all linked identities.
- › **Feedback loop:** fraud confirmed at redemption is used to retrain onboarding risk scoring, update device/account blacklists, and adjust in-survey and lifecycle rules.

What proves it works:

We couldn’t run the most demanding and visible parts of our business—from always-on brand tracking to high-stakes public election polling—without robust integrity. These use cases leave no room for unreliable respondents or weak safeguards.

Tracking: cleaner signals, steadier awareness baselines; reductions in low-quality completes without demographic drift.

Polling: high voting district level accuracy in national elections; model performance supported by verified respondents and deep respondent histories.

Operations: better raw data; live systems that adapt to new threat patterns.

At YouGov, integrity is built into the way we operate. Because we control the entire process—from recruitment, through every survey interaction, to the point a panelist redeems their rewards—we can apply the right techniques at the right time, informed by a unified view of risk across the panelist and project lifecycle. That control lets us connect signals over time, adapt to emerging threats, and act where it matters most, whether that’s blocking a bad actor at registration, flagging suspicious behavior mid-survey, or shutting down a payout attempt before money changes hands.

This holistic approach also means we don’t have to rely on heavy-handed measures at a single checkpoint. We can balance rigor with respect for genuine respondents, preserving the trust that keeps them engaged while removing those who don’t belong. By monitoring and improving our safeguards over time, we give our clients confidence that the people in our data are real, unique, and eligible.

At YouGov we regularly perform benchmarking exercises to ensure we are beating industry standards. We aren’t satisfied with just achieving better than average rates – we aim to deliver an under 2% fraud rate and regularly outperform this in the field.

Average fraud rate as high as

16%

Our commitment

<2%

How we performed in the benchmarking exercise

0.6%

*Based on analysis of a benchmarking exercise performed by YouGov across 12 global panel providers in November 2025

Accuracy

Accuracy asks whether the answers themselves are valid, in the sense that they are complete, thoughtful, and correct.

Accuracy is about ensuring that responses reflect the intent of the question and can be relied upon for decisions.



Beyond deliberate fraud an inattentive or disengaged human is also harmful to data quality. Sometimes the cause is the participant who, by rushing, straight-lining, skipping, or giving implausible answers, gives poor data. Sometimes the cause is the instrument, where ambiguous wording, long or complex questions, or onerous tasks can make it difficult for the participant to respond.

At YouGov, we address both sides. We design instruments that minimize burden and ambiguity, and we apply numerous in-survey specialized accuracy checks. This includes two signature methods: the Response Quality Score (RQS), a cross-panel metric built on hundreds of behavioral and technical indicators, and Awareness Cross-Entropy (ACE), a targeted technique for detecting anomalous brand awareness patterns in our daily tracking. These are integrated into operational processes so that low-quality data is removed before it can distort findings.

1. Designing for accuracy

Before the first answer is ever given, the quality of a dataset is shaped by the quality of its questionnaire. Poorly written, confusing, or overly long surveys create the conditions for disengagement and error, and no amount of post-collection cleaning can fully fix it. For this reason, we put as much effort into designing for accuracy as we do into measuring it.

Our aim is to make surveys easy to understand and natural to complete, regardless of device. This reduces respondent fatigue, eliminates avoidable confusion, and ensures that any quality problems we detect later are genuine rather than artefacts of bad design.

Our design principles include:

- › **Clarity and neutrality:** Wording is unambiguous and free from leading or loaded phrasing.
- › **Balanced answer sets:** Options reflect the full range of likely responses, with no implicit bias toward one end of a scale.
- › **Order control:** Randomizing the order of questions and answer options where appropriate to avoid position bias.
- › **Mobile optimization:** Layout and interaction design are tested to work well on all devices.
- › **Soft launches:** Pilots to catch comprehension issues, logic flaws, or excessive length before full deployment.

2. In-survey safeguards

Even with a well-designed questionnaire, participant engagement can vary. In-survey safeguards protect accuracy by identifying respondents whose answers are implausible or inconsistent as they are being given. This is where we can act immediately to stop bad data entering the set or flag a case for deeper review.

We combine direct checks, such as explicit attention prompts, with indirect ones that infer inattention or random clicking from behavioral patterns. Every safeguard is calibrated to the context so it can filter out genuine issues without penalizing careful respondents.

Our in-survey safeguards include:

- › **Attention checks:** Embedded instructions to confirm a respondent is reading carefully.
- › **Consistency checks:** Spotting contradictions within the same survey (e.g., claiming never to use a product, then describing it in detail) or across time (e.g., sudden changes in otherwise immutable attributes).
- › **Plausibility checks:** Identifying improbable claims, such as travel to highly unlikely destinations or holding contradictory extreme political opinions.
- › **Paradata analysis:** Monitoring device type, OS, page timings, and interaction patterns to detect rushing, straight-lining, or indiscriminate clicking.
- › **Open-end quality scoring:** Using a proprietary model to assess relevance and responsiveness; in higher-risk contexts, applying open-source LLM detection to identify AI-generated text.

3. Response Quality Score (RQS)

In-survey checks catch problems in the moment, but they can't see the bigger picture of a respondent's behavior over time. The Response Quality Score (RQS) was developed to fill that gap.

The RQS is a proprietary, respondent-level, longitudinal measure that detects patterns of low-quality behavior invisible to single-survey checks. This single score is derived from over 600 indicators of technical behavior, survey engagement, and content quality. We built it as a machine learning classifier and trained on human-coded examples of high- and low-quality respondents to tune its ability to catch genuine issues while avoiding false positives.

How it works:

We apply the RQS to monitor panel health and more deeply screen those panelists flagged by in-survey checks. RQS aggregates multiple types of data:

- › **Technical stability:** Device fingerprint consistency, geolocation stability, IP address patterns.
- › **Behavioral patterns:** Frequency and type of attention check failures, speeding thresholds exceeded, break-off rates.
- › **Content quality:** Open-end responsiveness, contradiction rates, prevalence of low-incidence claims.
- › **Profile integrity:** Changes in core demographics over time, consistency of declared vs observed attributes.

The RQS is the essential tool that informs our decisions to keep, quarantine, or remove respondents, and it feeds upstream into onboarding, adaptive in-survey checks, and broader integrity processes.

4. Awareness Cross-Entropy (ACE)

ACE is a YouGov-developed method for spotting unlikely or inconsistent patterns in brand awareness data—the kind of subtle anomalies that can slip through standard attention checks but still undermine the reliability of brand tracking. It uses a statistical measure called cross-entropy to compare each respondent's pattern of “heard of” and “not heard of” answers against the pattern we expect to see in the wider population. When the two differ sharply in implausible ways, ACE flags the case for removal or further review.

Daily brand tracking produces enormous volumes of awareness data, and those top-of-funnel figures underpin every other metric in the brand funnel. If awareness is wrong, consideration, usage, and advocacy measures will all be skewed. The challenge is that disengaged respondents don't always fail obvious quality checks, but their awareness answers may still be random enough to cause real damage to the data.

What is entropy?

Entropy, in information theory, measures the amount of unpredictability or disorder in a set of responses. If almost everyone says they've heard of Google, those answers have low entropy: they're predictable. If answers are all over the place for no apparent reason, entropy is high.

Cross-entropy compares two patterns of answers: the expected pattern in the population (based on historical data) and the observed pattern from one respondent. The more often a respondent disagrees with the majority in unlikely ways—for example, saying “no” to Google but “yes” to a niche local brand—the higher their cross-entropy score. High scores are a statistical red flag for implausible data.

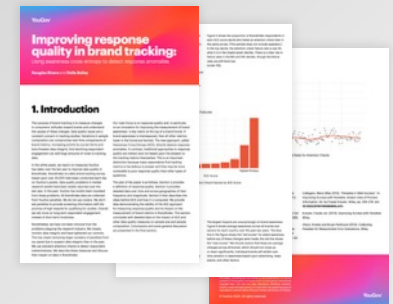
From concept to action

Once we've calculated ACE scores, we compare them to calibrated thresholds for each sector and country, because brand awareness norms differ by market and category. Respondents above the threshold are flagged, and their data is either removed or reviewed in conjunction with other indicators like RQS. This removes improbable swings in aided awareness, particularly for the most ubiquitous brands, and results in more stable month-on-month trends without demographic drift.

Want to know more?

Click here to read our whitepaper on ACE.

[Learn more](#)



Proven in Practice

All the techniques described above come together in a single discipline: deciding what to do with each interview so that only accurate, reliable responses enter the dataset. Sometimes that means including answers in full; sometimes it means tempering their influence, holding a case aside for further review, or removing it entirely when the evidence is clear. Because YouGov controls both the panel and fieldwork, these judgments can be made with full context, drawing on the respondent's history as well as their in-survey behavior.

By removing noise and preserving genuine market movement, our accuracy controls ensure that the underlying signal comes through without distortion. Trends remain clear and interpretable, so shifts in the data reflect what is happening in the real world rather than the by-products of poor response quality.

That clarity is what makes it possible to run the most demanding, high-visibility research with confidence, and it's why clients can rely on our outputs as a trusted foundation for decisions.



Improving accuracy in brand tracking

We have developed a proprietary metric we call Awareness Cross-Entropy (ACE), to address one of the most common problems that plague brand tracking studies: volatile data.

Over-time consistency is critical for brand tracking. For the most part, we expect awareness to be stable for most brands. Large swings in awareness are not plausible, unless brands are either rapidly growing, have massively increased their ad spend, or have been in the news—but these are the exception. A rise in awareness, followed by a decline, would imply that some people had forgotten about a brand that they recently knew about.

Entropy refers to the amount of disorder or uncertainty in a probability distribution. Cross-entropy compares the distance or discrepancy between two probability distributions. In this case, we are comparing the discrepancy between one respondent's answers and those given by a random draw from the rest of the population.

Our methodology works by identifying “outlier” behavior and applying a corrective factor to bring it closer to the norm. The net effect is significant dampening of unlikely volatility such that the real trends can come through.

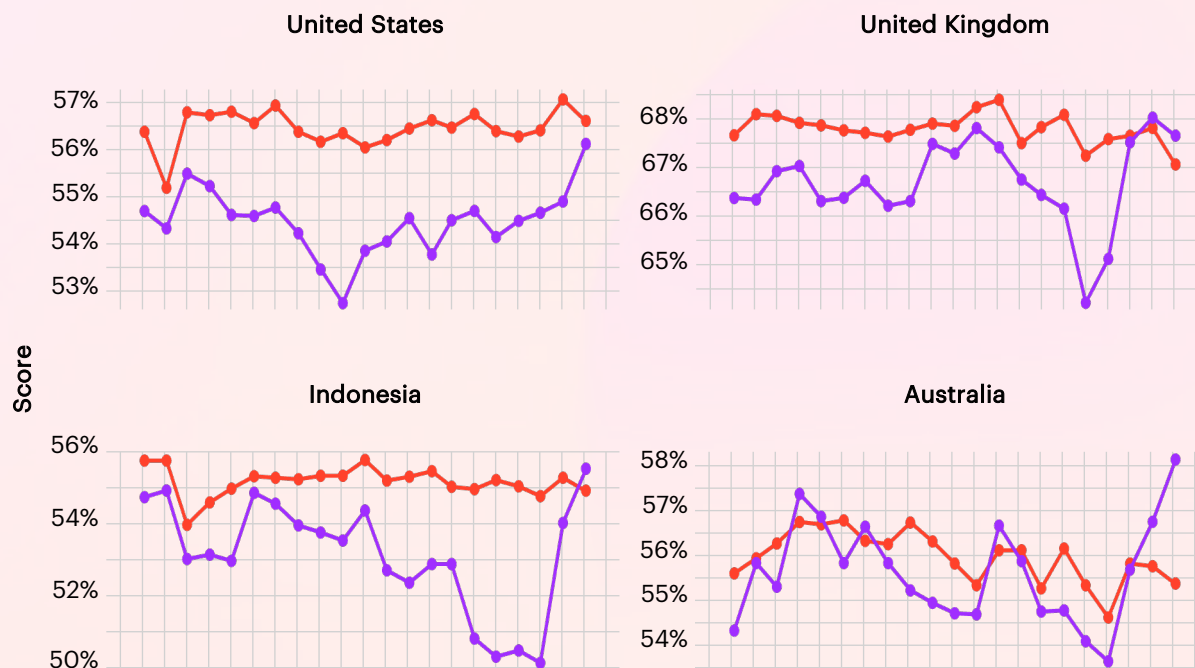
The charts on the right shows an example of the correction being applied. The line marked “Old” reflects the uncorrected data, while the line marked “New” reflects the data with the correction applied.

[Click here](#) to learn more about our Awareness Cross-Entropy Correction.

[Learn more](#)

Aided awareness

● New scores ● Old scores



Connected data

Connected data means that everything a panelist shares with us—who they are, what they think, and how they behave—is tied back to the same person through a single, lasting record.

That record grows over time. It starts with basic demographics and expands with each new survey answer, behavioral signal, or piece of external data a panelist chooses to share. Because all of this information is attached to the same person, different kinds of data—declared attitudes, observed behaviors, even passive or uploaded information—can be combined with confidence. This makes the pillars of quality stronger: we can check consistency across time, enrich a custom survey with existing profiles, or link brand tracking and polling back to the same people. For clients, it means the data they see is not a set of disconnected snapshots but part of a continuous, interpretable whole.

A longitudinal record

Connected data owes its existence to our ability to establish a durable, longitudinal record of a panelist's actions. From the moment someone joins the panel, we establish a record that combines their demographics, account details, and early quality checks. Every subsequent interaction—profile refreshes, daily brand responses, custom surveys, paradata, fraud controls and more—attaches to that same record. Over months and years of participation, this spine accumulates into a powerful longitudinal signal, showing not just what a person said at a point in time, but how their answers and behaviors evolve.

Multiple data streams

Three streams or types of data flow into the spine. Demographics anchor the record with facts like date of birth, region, and education. Declared attributes, refreshed through Profiles surveys, capture interests, attitudes, and category behaviors. Observed signals such as device type, timings, and response patterns add behavioral context. Some panelists also choose to share external data, such as financial records through openbanking which extend the view further. Because these are our panelists, not anonymous traffic, these streams can be combined confidently and reused across projects, improving targeting and accuracy while reducing participation burden on the panelist.

Strengthening the pillars

Connected data strengthens each of the quality pillars by turning what might otherwise be a single datapoint into a record that is broad, deep, and continuous.

Representation: A sample is never just a cross-section in time. Because panelists carry verified profiles that are refreshed regularly, YouGov can see whether the achieved sample continues to reflect the population. Longitudinal monitoring makes gaps visible and gives the confidence to screen eligibility from profiles rather than overburden surveys with screeners. The result is steadier representation and healthier panel engagement.

Integrity: Fraud is harder to hide in a connected environment. A duplicate account, a sudden location change, or an inconsistent claim is easier to detect when every response attaches to a persistent history. Because that history is cumulative, integrity isn't a one-shot check at sign-up, but a property that can be enforced throughout participation.

Accuracy: Accuracy is strengthened when answers can be cross-checked against what someone has said and done before. Declared behavior can be compared with observed signals, and attitudes tracked for plausibility over time. This continuity distinguishes genuine change from noise. Instead of relying only on in-survey checks, Connected data makes accuracy a longitudinal attribute of the panel.





Product implications

Connected data is not only a safeguard but also the foundation on which YouGov's products are built.

- › **BrandIndex** demonstrates its longitudinal power. Daily responses accumulate into stable, history-aware time series that capture real change while filtering out random noise.
- › **Polling** depends on connected data for accuracy. While every poll is cross-sectional in fieldwork, working with verified profiles and voting histories means the results can be modeled and predicted with confidence, even at the constituency and district level.
- › **Ratings** become powerful through connected data. Because they sit in the same Cube as Profiles, BrandIndex, polls, and custom studies, those simple signals can be tied to audiences and outcomes, making them useful for strategy and targeting.
- › **Profiles** shows its breadth. Hundreds of thousands of attributes can be linked to the same individuals whose opinions are tracked in BrandIndex or polled in elections, enabling precise segmentation and contextualization without burdening respondents.
- › **Custom Research** inherits the same advantages. A one-off study is enriched by existing profiles, targeted through known eligibility, and, when needed, extended through recontact.
- › **Crunch** makes connected data usable. It gives researchers a consistent cross-project audience view, linking BrandIndex, Profiles, and custom studies at the respondent level to enable analysis that would be impossible with disconnected datasets.

Because the architecture is single-source, it can be further extended. With panelist consent, we can link client databases, metered or passive data, and other external sources into the same respondent-level frame. This makes it possible to connect survey responses to actual behaviors, to enrich brand health metrics with transaction data, or to tie public opinion to media consumption—all within the same system.

Governance and trust

Connected data is credible only because safeguards are applied at every stage. Checks at onboarding, participation, and redemption are recorded against the same individual, so the trust in each panelist builds cumulatively over time. Privacy rules define what can be connected and how it can be used, and global governance ensures those standards are enforced consistently across markets. Put differently, connected data exists at scale because we treat the panel as a long-term relationship, not as a commodity.

Connected data is the foundation of everything YouGov does. It underpins every study, whatever the brief. It is a critical enabler of quality, and you can see it in practice every day in brand metrics that move with real market events, election forecasts that match actual returns, and custom surveys that can be extended and panelists recontacted without starting from scratch. Connected Data is the cumulative record of the demographics, beliefs, and actions of millions of individuals that make our data reliable, change interpretable, and insights actionable.

Connected data

YouGov connected data means that our data product tools are linked via single source.

Everything we see from a panelist—from profiling data to brand tracking data to qualitative data to a client's custom data and behavioral data feeds— can be used to develop rich, actionable insight for our clients.



Brand data

Brand tracking data on 27,000 brands across 55 markets spanning 17+ years of historical trends and performance benchmarks.



Profiling data

An ever-growing source of connected consumer data, with 2 million+ data variables from YouGov's 30 million+ global panel members. Covering demographic, psychographic, attitudinal and behavioral consumer metrics.



Custom data

Seamlessly connect your own custom data into YouGov's connected dataset.

Recontact highly specific audiences to dig deeper.



Behavioral data

Our technology links panelist's data (survey, social, behavioral) to their end-point behavior.

- > Media Consumption
- > Banking & Transactions
- > Online Search and Social
- > Gaming
- > Path to Purchase



3rd Party data

Link 3rd Party data sources into YouGov through various data base matching options.



Qual data

YouGov Qual data digs deeper into the drivers and motivations of the consumer mindset to add colour and further context to the breadth of connected data.

The panelist experience

From its founding, YouGov has been built on a simple conviction: the panel is the heart of the business. Everything—our accuracy, our credibility, our value to researchers and clients—depends on the people who choose to share their time and their opinions with us. The panelist experience is therefore crucial. From the moment someone joins the panel, their experience shapes whether they stay, how they participate, and the care they take with their answers.

YouGov's advantage lies in owning and managing our own panel. That control gives us the ability to shape the entire panelist journey—from the first moments of registration through to redemption and beyond. Each stage is designed with two goals in mind: safeguarding the quality of the data and sustaining the motivation of the people who provide it. When the experience feels smooth, fair, and respectful, panelists not only remain active but continue to contribute with care over time.

A virtuous circle

Recruitment is where the relationship begins. For the panelist, it's the moment they start judging us against our promises. It's also when we create an **identity spine**—a durable record that links demographics, declared attributes, and behaviors over time, and becomes the backbone of everything that follows. Every new panelist is welcomed with structured onboarding that validates who they are, confirms where they are, and sets clear expectations for participation. Fraudsters encounter friction, escalating checks, and ultimately removal; genuine people find the experience smooth and human. That asymmetry—rigorous for those who pose risk, seamless for those who don't—keeps the panel both protected and welcoming.

Profiling begins at the start, but it isn't a one-time exercise. We **enrich the identity spine** continually, adding new signals as people participate. The evolving record sharpens targeting, reduces repetitive questions, and lets us understand change rather than just collect snapshots. Because we also use profile data to improve the experience—shorter screeners, more relevant surveys—panelists stay engaged. Their continued participation, in turn, deepens the profile further, creating a cycle of knowledge that benefits both panelists and clients.

Sampling is where the promises of representation are put into practice. Because we own the panel, we decide who is invited, when, and how often. Each invitation is deliberate—designed to mirror the population, manage burden, and keep the panel healthy. Surveys are invite-only; there is no free-for-all or self-selection. Where possible, we use existing profiles to screen eligibility, shortening surveys and removing unnecessary frustration. Just as importantly, we manage cadence. We avoid overburdening in-demand groups, and we attempt to mitigate the hyperresponsive patterns often associated with so-called “professional respondents.”

Participation is the moment the relationship is tested. Every survey must balance the needs of research with the experience of the person answering it. We design studies for clarity and accessibility to minimize frustration and fatigue, enabling panelists to respond accurately. Background systems monitor behavior and responses before, during, and after fieldwork—through visible and invisible checks—so we can distinguish good respondents having an off day from bad actors. That fundamental fairness, coupled with attention to the user experience, is the only way to sustain engagement—and the only way to obtain continually accurate, representative data from real people.

Engagement at YouGov starts with a clear promise: if you're invited, there will always be a survey to take. That promise means panelists know they don't need to lie to qualify – they will get a survey anyway. This reduces screen-out frustration and builds trust from the outset. From there, panelists are kept in the loop through touchpoints like newsletters, micro-surveys, and feedback channels. These include notifications when their results appear in the news and Daily Questions where live results are visible in real time before final findings are published the next day. People tell us they join to have their voice heard and stay when they see their voices reflected back in our work.

Redemption closes the loop and is a proof point that we're keeping our side of the bargain. We pay quickly, transparently, and through flexible local options. 99% of valid (not risky) payments are made within 5 minutes of the request. Redemption is also our final checkpoint: we employ bank verification, sanctions screening, and device fingerprinting to deny fraudsters the rewards they seek.

Governance ensures that the panelist experience is more than an empty slogan. We embed our principles in the systems we build, enforce them through methods and processes, and refine them with evidence. Oversight keeps quality measurable and top of mind at the highest levels of our company. Metrics are reviewed regularly, panelist feedback is acted upon, and our practices align with international standards like ISO & TCF. Governance is the habit that ensures what we promise to panelists and clients is delivered consistently, at scale.





Why YouGov treats panelists as an asset, not a commodity

The practices underlying panelist experience yield more than just a better user journey. They embody YouGov's total commitment to data that reflects reality. Because panelists are respected, researchers can trust their answers. They are the fuel of our business and the reason YouGov can sustain daily brand tracking, call elections with precision, and provide evidence that clients can act on when the stakes are highest.

And behind it all is a mission we have carried from the start: to give people a voice. Panelists don't stay only for incentives; they stay because they want to see their voice reflected in public debate and in the decisions brands make. Initiatives like YouGov Plus extend that mission further, bringing trusted contributors into the process—testing new designs, critiquing survey experiences, and shaping our future direction. Their involvement is a reminder that panelists are more than respondents; they are partners.

In short, creating a great panelist experience is neither a box-ticking exercise nor a slogan. It is the keystone of our operation and the foundation on which representativeness, integrity, and accuracy rest.

Because the panel is treated as an enduring asset rather than a commodity, it continues to generate value for clients and meaning for panelists alike.

Our promise to panelists

Running research panels is a challenging task. That YouGov's panels are among the best in the business arises from six simple commitments to our members:

01

We keep our promises.

Every survey invitation leads to a live survey, not a long screening process. Members know their time is valued, and that reliability builds trust.

02

We show their impact.

Through our monthly YouGov in the News updates and You've Made the Headlines newsflashes, members see when their views appear in public debate.

03

We keep participation interesting.

Daily Questions and Daily Chat offer quick, engaging ways to share opinions, see instant results, and understand how their views fit into the wider world.

04

We recognize contribution.

You & YouGov gives members a personalized yearly summary, while their live Insights page keeps their statistics visible all year round.

05

We celebrate loyalty.

YouGov Plus recognizes our most active members with enhanced rewards, anniversary bonuses, and opportunities to shape what comes next.

06

We reward fairly.

Every contribution earns points that can be redeemed in ways that suit each member best, whether that's a bank transfer or gift cards for Amazon, ASOS, or Asda.

How YouGov uses external panel providers

While most YouGov research is conducted with our proprietary panel, we occasionally need to supplement with other sources, often to reach hard-to-find audiences. To do this, we work with trusted partners who specialize in certain groups and/or regions.

We take the partnership relationship seriously. Beginning from the moment we choose partners through fieldwork and into data processing and delivery, we aim to ensure that the partner's panelists provide the same level of quality as our own. Of course, we are transparent with our clients as well.

That said, every partner has its own way of doing things. Here is how we apply our lens of quality to the process, step-by-step.

Validation steps

Benchmarking

Our internal benchmarking studies play a critical role in how we choose external panel partners. By running the same tests on other providers that we apply to our own panel, we can see which suppliers deliver respondents who mirror census benchmarks, pass fraud and logic checks, and provide thoughtful, consistent answers. These studies give us a clear evidence base for partner selection. We work with those who meet our standards and avoid those who do not—ensuring that any third-party sample we blend into our studies strengthens, rather than undermines, the overall quality of the data. We always inform clients when we use external panels and will share anonymized data from these studies with them.

Partner panelists arrive at a YouGov survey

Partner panels are responsible for recruiting and preliminarily validating their panelists. Once a partner's respondent lands in a YouGov survey, our own onboarding defenses run automatically (e.g., device/geo risk signals), and we deduplicate across sources to prevent the same individual from entering via multiple routes. These safeguards are part of our standard setup for all interviews we conduct.

Profiling

YouGov panelists carry rich, continuously refreshed profiles that we can link over time; that's a property of our owned panel/Connected Data architecture. External respondents arrive with whatever attributes partners provide for eligibility and quota management, but they don't have the same persistent profile within our system and are not linked longitudinally in the same way. (We make this clear to clients.)

Participation & in-survey quality

Source does not change the checks we apply in-survey, and external samples go through far more checks. All respondents pass through our quality stack: third-party and proprietary signals (e.g., Research Defender, IP/geo/device indicators), paradata-based safeguards, and where applicable our specialized methods such as the Response Quality Survey (RQS) and Awareness Cross-Entropy (ACE) in daily brand tracking. These controls operate before delivery so low-quality or fraudulent cases are removed regardless of origin.

What this means for our quality pillars

Representativeness

External partners are brought in when we need to reach audiences that are rare, highly specific, or under-represented in our own panel. They extend coverage—for example, in small geographies, locations where we don't have a panel or niche target groups—but YouGov still manages the quotas and oversees how interviews are allocated. Because we control the frame and actively remove overlaps, the achieved sample continues to reflect the intended population rather than double-counting frequent survey takers or introducing hidden biases.

Integrity

Every respondent, whether from our own panel or a partner source, goes through onboarding and fraud checks.

Tools such as Research Defender are run universally, screening for duplicate devices, mismatched IP and geo signals, or suspicious behavioral patterns. These checks operate in real time, so even when respondents come from outside sources, they must clear the same barriers as our own panelists before their answers are accepted. Device fingerprinting, geolocation, and threat scoring mean that integrity is enforced at the interview level, not just at sign-up.

Accuracy

The safeguards that protect our panel also extend to partner-sourced interviews. Paradata such as timings and breakoff patterns are applied across all respondents. This means that poor-quality or implausible responses are filtered out before they ever enter the dataset, regardless of origin.

External panels complement YouGov's own panel when needed. They are used sparingly, for clearly defined purposes, and always under conditions that aim to maximize data quality. By applying strict controls—where possible, the same that we use on our proprietary panel—and by openly disclosing when and why we bring partners in, we ensure that additional reach does not come at the expense of trust.

Clients can be confident that even when partners are used, the resulting data meets the same standards of representativeness, integrity, and accuracy that define all YouGov research.

The reality of our research: proof of quality



The reality of our research: proof of quality

Quality is easy to promise and hard to prove. Lofty principles and corporate slogans aren't enough.

Transparency is part of the answer. The materials and data we publish go far beyond boilerplate or “me-too” statements meant to placate a casual reader or satisfy a procurement checklist. They reflect a sustained, tireless commitment to showing our work—openly and in detail—so others can see exactly how we deliver the outcomes we claim.

But the real proof is delivering representativeness, integrity, and accuracy consistently, at scale, in the most demanding settings. That's where YouGov stands apart.

Trust and accuracy

1. **The most trusted market research provider globally***
2. **Most quoted market research brand - trusted by the world media.****
3. **The market research provider most associated with delivering high quality, trusted data at scale.***

Our public polling record is one of the most visible, unforgiving tests of quality anywhere: election after election, our results have matched real-world outcomes across markets. Our daily brand tracking—thousands of interviews every day, sustained over years—demands stability and precision few organizations can match. And no agency is as widely cited in the global press. All these factors reflect the credibility of our data and the trust placed in it by journalists, analysts, decision-makers, and ordinary citizens.

The Best in the Business

Wondering how YouGov's polls performed in recent elections?

See our results from the UK 2024 elections, [the 2025 Australia Elections](#), [2025 state elections in the US](#) and read what [how the independent FiveThirtyEight website \(now ABCNews\)](#) ranked us one of the world's top pollsters.

While these are only recent results, YouGov has a track record of accurate polling dating back to our founding, beginning with the first elections we called: the 2001 UK General Election.

We mention this because our commitment to accuracy in polling is just as ever-present in our commercial work. We recognize that the stakes are just as high for our brand and retail clients. The work we do to ensure quality benefits **everyone** who relies on YouGov's data.

The results speak for themselves...

UK

Most accurate pollster with
93%
accuracy in 2024

Italy

Only
1%
error rate in 2023

US

Highest ranked
market research company methodology in 2024 (538)

Germany

92%
constituency accuracy in 2025

Spain

Correctly called
hung parliament in 2023

Australia

Only pollster
to call majority gov (97%)

* 3rd party piece of research across over 3000 market research users globally, Aug 2025

** Meltwater, 1 August 2024 to 31 July 2025

Quality isn't an aspiration at YouGov. It is the outcome of a philosophy that has guided us since our founding and the decisions we have taken over the years, from how we build our panel to how we design our surveys to how we hold ourselves accountable in public.

The proof lies not just in saying clients can trust the data, but in showing why they can: the transparency of our methods, the resilience of our systems, and the consistency of our results in showing the reality of public opinion.

That ability to connect principle, explanation, and evidence is what gives quality its meaning, and what allows us to say, confidently, to our clients:

**Yes, you can
trust the data.**



A photograph of three young people dancing outdoors under a clear blue sky. On the left, a young woman with short dark hair is wearing a green ribbed long-sleeved crop top and plaid shorts, with her arms crossed. In the center, a young man with short dark hair is wearing a blue and white striped short-sleeved shirt and teal shorts, with his mouth open in a joyful expression and arms raised. On the right, a young woman with long braids is wearing a bright pink strapless top and a purple skirt, with her arms raised behind her head. The YouGov logo is superimposed over the center of the image.

YouGov®

YouGov, 2026, all rights reserved. All materials contained herein are protected by copyright laws. Any storage, reproduction or distribution of such materials, in whole or in part, in any form without the prior written permission of YouGov is prohibited. This information (including any enclosures and attachments) is propriety and confidential and has been prepared for the exclusive use and benefit of the addressee(s) and solely for the purpose for which it is provided. We make no representations, warranties or guarantees, whether express or implied, that the information is accurate, complete or up to date. We exclude all implied conditions, warranties, representations or other terms that may apply and we will not be liable to you for any loss or damage, whether in contract, tort (including negligence), breach of statutory duty, or otherwise, even if foreseeable, arising under or in connection with use of or reliance on the information. We do not exclude or limit in any way our liability to you where it would be unlawful to do so.