



SHOT RESEARCH · 2026

THE GRASSROOTS CENSUS. **A METHODOLOGY FOR** **COUNTING SPORT** **HONESTLY.**

How SHOT proposes to measure grassroots sport participation from platform-recorded activity: what is counted, how privacy is protected, what will be published and when, and what the method cannot claim. This paper is a pre-commitment to a method. It reports no data, and it makes no claim that the data yet exists at publishable scale.

METHODOLOGY PAPER · VERSION 1.0 · JULY 2026 · FOR PUBLICATION AFTER INTERNAL
RATIFICATION OF THE MARKED PARAMETERS

SHOT Clubhouse Limited · Company no. 16350767 · 86-90 Paul Street, London EC2A 4NE

shotclubhouse.com

1 THE MEASUREMENT PROBLEM

Nobody can say with confidence how many people play grassroots sport, where, or how often. Every existing count answers a different question, and none of them answers that one.

REGISTRATION COUNTS THE AFFILIATED, NOT THE ACTIVE

Governing bodies count what their registration systems can see: members who have affiliated, paid a fee, or completed a form. This measures a transaction, not activity. A registered player who stopped attending in October is counted for the whole season; a player in an unaffiliated five-a-side league, a casual weekend game, or an informal school-adjacent session is not counted at all. The unaffiliated majority of recreational sport is structurally invisible to registration-based counting, however carefully the registers are kept.

In conversations with governing bodies and funders during 2025-26, SHOT has repeatedly heard of sports where the registered figure and credible estimates of actual players diverge by more than an order of magnitude. These accounts are reported here as experience, not as verified statistics, and no body is identified.

SURVEYS SAMPLE, AND THIN OUT FAST

Survey measurement is currently done best by Sport England's Active Lives surveys, which ask large national samples about their activity and publish on consistent definitions with a long time series. Active Lives is the standard this paper treats as current best practice, and its limits are the limits of any survey. Responses are self-reported, subject to recall error and social desirability bias. The sample, though large nationally, thins quickly below national level: small sports, small areas and small demographic groups produce wide confidence intervals or no publishable estimate. The instrument observes what people say they did over a reference period, not sessions as they happened, and its cadence is measured in months.

BODIES MEASURE INCOMPATIBLY

Even where counts exist, they do not add up, because each body defines participation differently: once in the last month, twice in the last 28 days, weekly during the season, ever in the last year. Age bands differ, seasons differ, and the same person appears in several registers or none. Figures produced under incompatible definitions cannot be summed or compared, yet funding decisions are routinely made by comparing them.

The result is a sector that allocates money against numbers no two parties define the same way. The remedy is not another estimate. It is a stated method: what is counted, how, and under what protections, published before any figure is.

2 WHAT THE PLATFORM OBSERVES

SHOT does not run a survey. The platform records participation events as a by-product of clubs running their week. The census is built, and only built, from those events.

DEFINITIONS

The following terms are used throughout this paper and will be used, unchanged, in any published census. Changes to these definitions follow the protocol in Section 5.

- **Club.** An organisation (club, team, school or programme) that administers sessions on the platform.
- **Session.** A scheduled instance of sporting activity created by a club or coach, carrying a date, a start time, a sport, an age group and a location area. The location is held at area level for census purposes, not as a precise venue coordinate.
- **Attendance event.** A record that a specific registered participant was present at a specific session, entered by the coach or organiser at or near the time of the session.
- **Evaluation event.** A structured development rating recorded within a coaching framework, by a coach, by the participant, or both. Evaluation content is never published; the census uses only its existence and timing.
- **Active participant.** A registered individual with at least one attendance event in the reference period. The proposed reference period is a rolling 28 days; this parameter is marked for ratification in Section 5.
- **Retention indicator.** The proportion of active participants in one reference period who remain active in a later period, computed only on consenting records.

PROPERTIES OF THE DATA

Three properties distinguish this record from both registration and survey data. It is behavioural: an attendance event describes what happened, not what someone recalls. It is session-level: activity is observed per occasion, not summarised per year. It is longitudinal: the same consenting individual can be observed across weeks and seasons, which is what makes retention measurable rather than inferable.

One property must be stated with equal weight. This is a register of activity on one platform, generated by the clubs that choose to use it. It is not a sample of the population, and Section 6 sets out what follows from that.

3 THE PRIVACY MODEL

The census exists only if the people counted control their part in it. The privacy model is therefore prior to the statistics, not an appendix to them.

OPT-IN BY DESIGN

No individual's records enter census aggregates without explicit, informed consent to research use. That consent is collected separately from the consent needed to use the platform, is refusable without any loss of platform functionality, and is withdrawable at any time. Withdrawal removes the individual's records from all future aggregation runs. For participants under 16, census consent is given by a parent or guardian; for the youngest tier it is given and managed entirely by the guardian.

DATA MINIMISATION AT REGISTRATION

The platform collects at registration only what is needed to run sport safely: a name, a date of birth (required for age-tiering and safeguarding), and contact details for the account holder. Demographic fields used in census breakdowns, such as gender and disability status, are optional, clearly marked as optional, and collected with their census purpose stated. No field is collected because it would be statistically interesting; every field must justify itself operationally first.

AGE-TIERED ACCOUNTS AND GUARDIAN CONSENT

Accounts are age-tiered at under-13, 13-15 and 16 and over, with guardian consent and guardian visibility built into the structure for minors. This mirrors the platform's safeguarding architecture and is designed to conform to the ICO's Age Appropriate Design Code. The census inherits these tiers: age is only ever published in bands, never as a date of birth or exact age.

THE INDIVIDUAL OWNS THE RECORD

Each participant, or their guardian, owns their participation record. They can export it in a machine-readable form and take it with them, and they can delete it. Deletion removes the record from all future aggregates. Previously published aggregates are not withdrawn, because by construction they contain nothing attributable to any individual.

No personally identifiable information crosses the aggregation boundary. No name, date of birth, contact detail, precise location or free text ever appears in, or is used to compute, a published figure. Published outputs contain counts and rates only.

4 AGGREGATION RULES

The census applies standard statistical disclosure control, stated in advance, so that no individual and no single club is ever inferable from a published figure.

MINIMUM CELL SIZES

No published figure describes fewer than a minimum number of individuals. The proposed minimum cell size is ten consenting individuals per cell. Any cell below the floor is suppressed, and complementary suppression is applied so that a suppressed cell cannot be recovered by subtraction from published totals.

GEOGRAPHIC AND DEMOGRAPHIC FLOORS

- **Geography.** Nothing is published below local authority level, and a local authority figure is published only where it also clears the cell-size floor. Club-level figures are never published; a club's own data is visible only to that club.
- **Demographics.** Cross-tabulations (for example sport by area by age band by gender) are published only where every cell in the table clears the floor. Where a breakdown fails, categories are aggregated upward until it passes, or the breakdown is dropped.
- **Age.** Published only in the platform's age bands, never in single years of age.

DOMINANCE AND ROUNDING

Where a single club contributes more than a stated share of a cell (proposed: half), the cell is suppressed or aggregated upward, because the figure would in effect describe one organisation. All published counts are rounded (proposed: to the nearest ten) as an additional barrier against differencing attacks across releases.

THE RELEASE TEST

Before any release, every table is tested against one question: could any individual or any single club be identified or inferred from this output, alone or in combination with publicly available information? If the answer is anything other than no, the output is aggregated upward or withheld. This test is applied by a named reviewer and recorded for each release.

PROPOSED PARAMETERS, FOR INTERNAL RATIFICATION: minimum cell size 10; geographic floor at local authority level; dominance threshold 50% from a single club; rounding to the nearest 10. These values follow common practice in official statistics disclosure control and will be ratified, and published as ratified, before the first census release.

5 PUBLICATION PROTOCOL

What will be published, at what cadence, under what caveats, and, critically, not before the thresholds stated here are met.

WHAT WILL BE PUBLISHED

An annual Grassroots Census report containing: counts of active participants and sessions delivered; attendance rates; and retention indicators; broken down by sport, region, age band and gender wherever Section 4 permits. Every edition carries a methodology annex restating the definitions in Section 2, the parameters in Section 4, and the coverage statement required by Section 6. Interim topic notes may be published between editions under exactly the same rules.

CAVEATS ATTACHED TO EVERY FIGURE

No census figure will be presented as a national estimate of sport participation. Every figure describes consenting activity recorded on the platform, and says so where it is printed, not only in a footnote. Comparisons with national surveys are presented as calibration, as described in Section 6, never as correction of them.

METHODOLOGY CHANGES ARE PUBLISHED BEFORE USE

The methodology is versioned. Any change to a definition, parameter or rule is published before it is used in a release. Where a change breaks comparability, the back series is either restated under the new method or marked with an explicit series break. Errors discovered after publication are corrected and the correction noted; nothing is amended silently.

PUBLICATION THRESHOLDS

The census will not be published until the consenting base is large enough to mean something. The thresholds below are placeholders, stated now so they cannot be quietly lowered later, and marked for internal ratification. No current platform counts are reported in this paper, and none of these thresholds should be read as a claim about them.

- **Sport-level figures** (proposed): at least 50 clubs and at least 1,000 consenting active participants in that sport.
- **Regional figures** (proposed): at least 20 clubs in that region, in addition to all Section 4 floors.
- **First edition** (proposed): published only when at least three sports and at least three regions independently clear their thresholds.

Commitment: if the thresholds are not met, the census is not published. A late, honest census is worth more than an early, thin one.

6 LIMITATIONS

A method is only as honest as its statement of what it cannot do. These limits will be restated in every census edition.

- **Self-selection.** Clubs choose to use the platform, and participants choose to consent. Both choices are non-random, so the census is not a probability sample and no weighting scheme can fully repair that. Platform-using clubs may differ systematically from others in size, organisation and digital confidence.
- **Coverage stated, every time.** Each edition will state the number of clubs, sports and regions represented and, as plainly, the populations it cannot see: informal and unorganised sport, non-consenting participants, and clubs not on the platform. The census under-counts informal sport by construction.
- **Recording quality.** Attendance is entered by coaches and organisers, and diligence varies. Data quality checks will be described in the methodology annex, and known weaknesses reported rather than smoothed over.
- **Digital exclusion.** A platform-based census risks missing exactly the communities that other counting also misses. This is a limitation to be measured and narrowed, not an inconvenience to be footnoted.
- **Calibration, not competition.** National surveys, with Active Lives foremost, remain the population-level estimate of participation. The census adds session-level behavioural depth that surveys cannot reach. Where census trends and survey trends diverge, the divergence will be published and examined, not arbitrated in the census's favour.

7 GOVERNANCE

The methodology is reviewed annually, and the review is published whether or not anything changes. External advice is welcomed: statisticians, safeguarding specialists, governing bodies and funders are invited to critique this method before the first release, and a named contact for methodological challenge will be published with every edition. Although SHOT is not a producer of official statistics, the protocol is written to the spirit of the UK Statistics Authority's Code of Practice for Statistics: trustworthiness, quality and value, in that order.

This paper commits SHOT to a way of counting before there is anything to count at publishable scale. That is the point. The method is only credible if it was fixed before the results were known.

SOURCES AND STANDARDS

This is a methodology paper. It reports no participation data, and the sources below are cited for methods, standards and context, not results.

SURVEY BENCHMARKS

- Sport England, Active Lives Adult Survey: methodology and publications (ongoing series)
- Sport England, Active Lives Children and Young People Survey 2024-25 (published December 2025)

STATISTICAL STANDARDS AND DISCLOSURE CONTROL

- UK Statistics Authority, Code of Practice for Statistics (trustworthiness, quality, value)
- Government Statistical Service, guidance on statistical disclosure control for tables and microdata
- Information Commissioner's Office, guidance on anonymisation and pseudonymisation

PRIVACY AND CHILDREN'S DATA

- UK GDPR and Data Protection Act 2018: lawful basis, consent and data minimisation principles
- Information Commissioner's Office, Age Appropriate Design Code (the Children's Code)
- Ofcom / UK Government, Online Safety Act age-assurance duties, enforceable from July 2025

CONTEXT

- Accounts of divergence between registered and actual playing populations are drawn from SHOT's own conversations with governing bodies and funders during 2025-26. They are reported in Section 1 as experience, are unverified, and identify no organisation.

Version 1.0, July 2026. Parameters marked as proposed in Sections 2, 4 and 5 require internal ratification before the first census release; the ratified values will be published in this paper's next version before any data is. Comments and methodological challenge are welcome via shotclubhouse.com.