



SHOT RESEARCH · 2026

THE CONFIDENCE DELTA.

A pre-registered methodology for testing the self-versus-coach rating gap as a retention signal in youth sport. This document fixes our hypotheses, definitions, thresholds and analysis plan before the data exists, so that whatever we publish later can be trusted.

PRE-REGISTRATION · VERSION 1.0 · JULY 2026 · NO RESULTS REPORTED

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

shotclubhouse.com

1 BACKGROUND AND HYPOTHESIS

Adolescence is where sport quietly loses children, and confidence is the word that keeps appearing in the evidence. This study pre-registers a test of whether a measurable confidence signal can warn of drop-out before the attendance register does.



THE INSTRUMENT

In PERFORM, SHOT's development engine, an athlete self-rates their confidence before and after each session on a 1-5 development ladder. The coach rates the same athlete independently on the same ladder, and neither sees the other's rating at entry. Every rated session therefore produces a paired observation. The difference between the two, self-rating minus coach rating, is the rating gap.

WHY THE GAP MIGHT MATTER

The disengagement literature points at self-perception, not ability: girls who leave sport cite low confidence and fear of judgement far more often than lack of interest. The parental literature points the same way. Peer-reviewed studies consistently find that perceived pressure predicts drop-out, anxiety and lower enjoyment, while perceived support predicts enjoyment, motivation and retention. A child who persistently rates herself below her coach's independent view may be carrying exactly the self-doubt this literature describes, in a form the register cannot see. Whether that signal predicts anything is the question this study exists to test, not to assume.

HYPOTHESES

1. **H1.** A persistent negative rating gap, the athlete rating themselves below the coach, predicts elevated drop-out risk within the following two blocks (eight weeks).
2. **H2.** The gap widens before attendance decays. It is a leading indicator of disengagement, not a lagging one.
3. **H3.** The effects in H1 and H2 are stronger for girls aged 11-15, consistent with the disengagement literature.

Each hypothesis is stated so it can fail. The definitions that make them testable follow in Section 3.

2 WHY PRE-REGISTER

A finding is only worth as much as the discipline behind it. Pre-registration means the methods are fixed and published before the data can influence them, so the results cannot be quietly shaped to fit a story.

The failure mode we are guarding against is well documented: run many analyses, keep the one that flatters, and present it as the plan all along. A platform testing its own feature has an obvious interest in a positive result, which is precisely why the test must be designed so that we cannot give ourselves one. This document is that design.

1 HYPOTHESES FIXED IN ADVANCE

H1, H2 and H3 are the only confirmatory hypotheses. Anything else we notice in the data is labelled exploratory, published as such, and never promoted after the fact.

2 DEFINITIONS FROZEN

The rating gap, drop-out, and every covariate are operationalised in Section 3 before analysis. No definition changes once outcome data has been seen.

3 THRESHOLDS STATED BEFORE THE DATA CLEARS THEM

Section 4 sets the minimum dataset required for publication. We publish when the data crosses the line, not when the result looks good.

4 EVERY CHANGE LOGGED

Any amendment is recorded in a version log with the date, the change and the reason, and the log is published with the results. Silent revisions do not happen.

5 NULL RESULTS PUBLISHED

If the rating gap predicts nothing beyond attendance alone, that finding is published in full under this pre-registration, with the same prominence a positive finding would receive.

Results will be published only against the methods stated here. If the methods change, the change and its reason are published too. That is the whole deal.

3 VARIABLES AND DEFINITIONS

A hypothesis is only testable once its terms are pinned down. These are the operational definitions this study will use.

THE EXPOSURE: THE RATING GAP

For each rated session, the rating gap is the athlete's post-session confidence self-rating minus the coach's independent rating of the same athlete, on the same 1-5 development ladder. A negative gap means the athlete places themselves below the coach's view. The pre-session self-rating is retained for sensitivity analysis but is not the primary measure.

A **persistent negative gap**, the exposure in H1, is defined as: a mean gap of minus 0.5 or lower across the athlete's most recent six rated sessions, with at least four of those six individual gaps negative. Rolling windows of four and eight sessions will be run as pre-specified sensitivity analyses; the six-session window is primary.

THE OUTCOME: DROP-OUT

Drop-out is defined as no attendance at any scheduled session for **six consecutive weeks** in which the athlete's group had at least one session scheduled, following prior regular attendance, meaning at least 60 per cent of scheduled sessions across the preceding two blocks (eight weeks). Weeks with no scheduled sessions, such as season breaks, do not count towards the six; the clock pauses and resumes. An athlete who transfers to another team or club within the platform, with continuity of record, is not a drop-out.

COVARIATES

- **Age band:** under 11, 11-15, 16 and over, from date of birth held under the account's consent regime.
- **Recorded gender,** where consent covers its use in research; otherwise excluded from the H3 subgroup, never imputed.
- **Sport and framework level,** one of the seven PERFORM pathway levels.
- **Attendance history:** attendance rate and trend over the prior two blocks.
- **Tenure:** time since the athlete's first rated session on the platform.
- **Season block,** to absorb the timing effects of the sporting calendar.

TIME ZERO AND CENSORING

An athlete enters the risk set at the first rated session after accruing the qualifying attendance history. Observation ends at drop-out, at the close of the study window, or at loss of record continuity, whichever comes first. Athletes still attending at the close of the window are censored, which the analysis method in Section 5 handles without discarding them.

4 DATA THRESHOLDS FOR PUBLICATION

We will not publish on a dataset too small to mean anything. The thresholds below are the proposed minimums the data must clear before any results paper is written.

2,000

athletes with a qualifying attendance history and at least one full block of paired ratings

Proposed minimum

150

observed drop-out events across the analysis population

Proposed minimum

8

sports represented, spanning at least three of the seven pathway levels

Proposed minimum

26

weeks of observation, with median follow-up of at least two full blocks

Proposed minimum

The event count matters as much as the athlete count. In time-to-event work, statistical power comes from the number of drop-outs observed, not the number of people watched, so a large panel with few events would still fail the bar.

Two further conditions apply. First, paired-rating completeness: at least 60 per cent of attended sessions must carry both a self-rating and a coach rating, so the exposure is measured rather than guessed. Second, subgroup protection: no subgroup finding, including H3, will be published from fewer than 300 athletes and 30 events. If the girls-aged-11-15 subgroup falls short, H3 is reported as untestable, not stretched.

Status: these thresholds are proposed and marked for internal ratification before data collection completes. Any change made at ratification will be recorded in the version log and published, per Section 2.

If the population grows more slowly than expected, the correct response is to wait, not to lower the bar. There is no publication date; there is a publication threshold.

5 ANALYSIS PLAN

The core question is not "do gap-carriers drop out more" but "do they drop out sooner, and would we have known anyway from attendance alone".

TIME-TO-EVENT FRAMING

Every athlete in the study has a clock that starts at their entry to the risk set. We follow each athlete week by week and ask whether those carrying a persistent negative gap leave sooner than comparable athletes who do not. This is survival analysis, the standard method for time-to-event questions. Its advantage over a simple percentage comparison is honesty: athletes observed for different lengths of time, and athletes who never drop out at all, are counted correctly rather than discarded.

The primary model is a proportional hazards regression with the persistent-gap flag as a time-varying exposure, adjusted for the covariates in Section 3. In plain terms: at each point in time, does carrying the flag raise the risk of leaving within the next two blocks, holding age band, sport, level and attendance history constant. Effects will be reported with uncertainty intervals, not bare headline numbers.

THE BASELINE COMPARISON

H1 only earns its keep if the gap tells us something attendance could not. The comparison model uses attendance alone: rate and trend over the prior two blocks, plus the same covariates, minus any rating data. We then test whether adding the rating gap improves prediction, judged two ways: discrimination, meaning how often the fuller model correctly ranks the athlete who left as higher risk than the one who stayed, and a formal likelihood comparison between the two models.

THE LEADING-INDICATOR TEST (H2)

Among athletes who drop out, we chart the rating gap against attendance over the eight weeks before exit. The pre-registered test is whether the persistent-gap flag fires earlier than an equivalent attendance-decline flag, with the median lead time and its range reported. If the gap moves with attendance rather than ahead of it, H2 fails.

SUBGROUPS, CORRECTIONS AND MISSINGNESS

H3 is tested as an interaction term for girls aged 11-15, subject to the subgroup thresholds in Section 4. Because three hypotheses give three chances of a fluke, significance levels carry a standard sequential correction for multiple testing, specified in the analysis code before unblinding. Missing self-ratings are treated with care: a disengaging athlete may stop rating before they stop attending, so a pre-specified sensitivity analysis treats missing self-ratings as a signal in their own right rather than a hole in the data.

Falsification condition: if the rating gap adds no predictive power over attendance alone, that result will be published in full under this pre-registration. A clean null is a publishable finding of this study, not a failure of it.

6 ETHICS AND PRIVACY

This is a study about children's confidence, run by the platform they use. The privacy bar is therefore set higher than the legal minimum, and it is structural, not procedural.

ANONYMISED ANALYSIS

Analysis is performed on pseudonymised records, with direct identifiers removed before any researcher touches the data. Published outputs are aggregate only. No individual child, coach, team or club will be identifiable in anything this study releases, and no club-level or team-level results will be published.

CONSENT REGIME

Under-16 accounts on the platform exist under guardian consent. The analysis population is restricted to accounts whose consent covers the use of activity and rating data for research in anonymised, aggregated form. Where consent does not cover a field, such as recorded gender, that field is excluded rather than inferred. Withdrawal of consent removes the athlete's data from all future analysis runs.

SMALL-CELL PROTECTION

No published figure will be drawn from a cell of fewer than ten athletes. Any row, subgroup or breakdown below that line is suppressed or merged before publication, so no combination of published aggregates can single out a child.

THE STUDY IS ANALYSIS, NOT INTERVENTION

Nothing in this study labels a child. No risk flag is shown to athletes, parents or coaches as part of the research; ratings remain what they already are on the platform, a coaching and reflection tool. If the signal proves useful, any product use of it would be designed separately, with its own safeguarding review. This pre-registration covers the test, not a deployment.

The study design and this document pass through internal ethics review before data collection completes, alongside the threshold ratification in Section 4; both reviews are recorded in the version log.

The child is the point of the study and the constraint on it. Nothing we learn is worth a single identifiable child in a published table.

7 LIMITATIONS

Stating the weaknesses now is part of the discipline. These are the limits any results paper under this pre-registration will carry on its face.

- **Observational, not causal.** If the gap predicts drop-out, that does not show the gap causes it, nor that closing it would keep a child in sport. This study tests a signal, not a treatment.
- **Not a probability sample.** The platform population is drawn from clubs that chose to adopt PERFORM, not randomly from grassroots sport. Generalisation beyond comparable settings will be claimed cautiously or not at all.
- **Measurement noise.** Children's self-ratings are imperfect, coaches calibrate the same 1-5 ladder differently, and a five-point scale is coarse. Noise of this kind usually weakens signals rather than inventing them, but it is noise all the same.
- **Informative missingness.** A disengaging athlete may stop rating before they stop attending, which could exaggerate the apparent lead time in H2. The sensitivity analysis in Section 5 addresses this but cannot eliminate it.
- **Drop-out from the record is not drop-out from sport.** An athlete who leaves a platform club may continue playing elsewhere, unseen. Our outcome measures departure from the observed setting, and the results will say so.
- **Subgroup power.** H3 depends on a subgroup that may not clear the Section 4 thresholds within the first window, in which case it is reported as untestable rather than answered thinly.

SOURCES

- Women in Sport, survey of more than 4,000 teenage girls: 43% of "sporty" girls disengage after primary school against 24% of boys; 68% cite fear of judgement and 61% low confidence
- Youth Sport Trust, Girls Active survey, 2024: girls' enjoyment of PE falls from 86% at ages 7-8 to 56% at ages 14-15
- Sport England, Active Lives Children and Young People Survey 2024-25 (published December 2025): "sport matters to me" falls from 51% to 39% between primary and secondary; 49.1% of children meet the daily activity guideline
- World Health Organization, 2019: 81% of adolescents aged 11-17 globally are insufficiently active; 84.7% of girls
- Peer-reviewed studies of parental involvement in youth sport: perceived pressure predicts drop-out, anxiety and lower enjoyment; perceived support predicts enjoyment, motivation and retention. Sport, Business and Management (2025/26); a youth football study indexed in PubMed Central (2021); a phenomenological study in an MDPI journal
- Scoping review of coach-parent communication, International Review of Sport and Exercise Psychology, 2024: regular, honest communication builds trust and improves both coaching and parenting
- Study of parent-coach communication frequency and satisfaction, International Journal of Sport and Exercise Psychology, 2025 (cited for framing; full text paywalled)

This is a methods document. It contains no results and no SHOT platform data. The statistics quoted in Section 1 are external context from the sources above, cited at claim level. Version 1.0, July 2026; all amendments will appear in the published version log.