

Tucson State of Reading

2026 BASELINE EDITION

Wave 1 baseline frozen

Methodology Documentation

THE TECHNICAL COMPANION OF THE 2026 BASELINE EDITION

Produced by The Place & Page
Tucson, Arizona

ABOUT THIS DOCUMENT

Document status

Scoring architecture and component weights locked June 2026. Survey instrument and fielded routing specified July 2026. Wave 1 analytic record finalized 17 August 2026. The dated decision record appears in Section 11.

This methodology preserves the five-component TRHI™ architecture and component weights while reconciling the fielded routing, path-specific scoring, final weighting, sample disclosure, open-end governance, qualitative coding, and limitations to the frozen Wave 1 analytic record. Where an item could not be traced to a saved source, it is stated as unavailable rather than reconstructed.

Wave 1 at a glance

WAVE 1 FACT	FINAL VALUE
Qualified completes delivered	520
Excluded after QC/geography review	9
Final analytic sample	511
Path A — chosen reader	431 (84.3% unweighted)
Path B — required-only reader	48 (9.4% unweighted)
Path C — non-reader	32 (6.3% unweighted)
Final weighted population TRHI	39.70 — Emerging (interpretive range)
Final weighted Path A TRHI	45.44

Figures are the frozen Wave 1 analytic record. Path shares are unweighted; the index values are weighted.

Contents

1	Study Overview and Purpose	5
2	Geographic Scope and Population Definition	6
3	Survey Instrument	7
4	TRHI Scoring Methodology	9
5	Bilingual Instrument and Translation	15
6	Sample and Panel Methodology	16
7	Data Quality Control and Analytic Governance	19
8	Post-Stratification Weighting	22
9	Theoretical Framework and Civic Purpose	25
10	Known Limitations and Disclosure Framework	26
11	Methodology Decisions — Formal Changelog	30
A	Appendix A — Final Weighting Benchmarks and Diagnostics	32
B	Appendix B — Final Wave 1 TRHI Score Summary	34
C	Appendix C — Public/Private Data Firewall and Open-End Coding Governance	35
	References and Source Notes	37

REFERENCE

Tables

LIST OF TABLES

Table 1	Reader paths and gateway definitions	7
Table 2	Component architecture and weights	9
Table 3	Reading frequency scoring	10
Table 4	Reader identity scoring	10
Table 5	Social reading scoring	11
Table 6	Path-specific structural scoring	12
Table 7	Path C structural-zero sensitivity	12
Table 8	Score bands and publication treatment	13
Table 9	Weighted and unweighted comparison	14
Table 10	Final analytic sample by path	17
Table 11	Raking margins and ACS sources	22
Table 12	Calibration balance	23
Table 13	Final weight diagnostics	24
Table 14	Methodology decision record	31
Table A1	Age benchmarks	32
Table A2	Education benchmarks	32
Table A3	Hispanic origin benchmarks	32
Table A4	Weight diagnostics	33
Table B1	Wave 1 TRHI score summary	34

Study Overview and Purpose

The Tucson State of Reading™ (TSOR™) is an independent annual civic research initiative measuring the health of voluntary reading culture across Pima County, Arizona. TSOR is produced by The Place & Page, a forthcoming independent bookstore in Tucson.

TSOR's primary output is the Tucson Reading Health Index™ (TRHI), a composite 0–100 index designed to measure the breadth, depth, and civic embeddedness of voluntary reading culture across the adult population of Pima County. The index is intended as an annual benchmark that can be repeated using the same instrument, scoring architecture, and core analytic rules.

TSOR is not a literacy assessment. It does not test reading skill or proficiency. It measures reading as a voluntary cultural practice: whether adults choose to read or listen, how reading relates to their lives, how often reading is shared socially, and how people participate in the cultural and civic ecosystem surrounding reading.

Wave 1, fielded in summer 2026, establishes the baseline against which future waves will be compared. The value of the baseline is longitudinal: later editions can ask what changed, what remained stable, and which components moved in different directions without redefining the index each year.

Geographic Scope and Population Definition

2.1 Target Population

The TSOR target population is adults age 18 and older residing in Pima County, Arizona. The adult population denominator used for Wave 1 weighting is 852,768 adults age 18+, from the U.S. Census Bureau American Community Survey (ACS) 2024 5-Year Estimates, Table B01001.

- IMLS POPU_LSA is not used as the county adult denominator because it is a library-service-area construct and does not represent Pima County adults.
- IMLS POPU_UND is used only for library per-capita measures where that denominator is appropriate; it is not the TSOR population denominator.

2.2 Geographic Branding vs. Measurement Unit

The Tucson metropolitan area, as the federal government defines it, is Pima County — a single-county metropolitan statistical area. The two names describe the same place. TSOR uses Tucson when referring to the community and Pima County when referring to the measured population. Pima County includes the City of Tucson and surrounding incorporated and unincorporated communities, and every public statistical claim must identify Pima County adults as the measured population when geographic precision matters.

Publication wording: “The Tucson Reading Health Index is calculated from a weighted online-panel sample of Pima County adults. Pima County is the geographic unit of measurement; the Tucson name is the civic brand for the regional reading-culture initiative.”

Survey Instrument

3.1 Architecture and Authoritative Fielded Source

Wave 1 used a branching survey administered in English and Spanish through Cint/Forsta. The authoritative source for actual fielded routing and response options is the final Cint / Forsta Survey Programming Specification. Earlier instrument drafts remain part of the design history, but the programming specification controls whenever wording or routing conflicts with an earlier document.

The instrument contains: (1) public P-series questions measuring reading behavior, access, identity/relationship, social participation, civic perception, and an open-ended reading-city question; (2) D-series demographic questions; and (3) a separate T-series module, not used in public analysis. The T-series is never used to produce or editorially shape public TSOR findings.

3.2 Path Structure

PATH	GATEWAY DEFINITION	FIELDIED PUBLIC-QUESTION PATTERN
Path A — Chosen Reader	At least some reading/listening is by choice	Receives full voluntary-reading path, including P1, P10, P11 and P12.
Path B — Required-Only Reader	Reads only what work or school requires	Does not receive P1, P10, or P11; receives P2/PP/P3/P3A/P4/P5/P7/P9/P12/P13/P14A/P14.
Path C — Non-Reader	Does not read/listen by choice or otherwise	Receives P3/P3A/P4/P7/P13/P14A/P14; does not receive P1/P2/PP/P5/P9/P10/P11/P12.

Table 1. Reader paths and gateway definitions.

P0 is the gateway. The paths are mutually exclusive and exhaustive. Path assignment is based on kind of reading (chosen, required-only, or none), not on the amount of reading.

3.3 Survey Preamble

All respondents were told that the study asks Pima County adults about reading and literary culture “whether you read a lot, a little, or not at all,” that there are no right or wrong answers, and that all experiences are valuable. This framing was intended to reduce self-screening by non-readers before P0 assigned the respondent path.

3.4 Key Design Decisions — Final Wave 1 Record

- P0 separates chosen reading, required-only reading, and no reading/listening so the population index does not condition only on existing voluntary readers.
- P1 measures time spent on voluntary reading and is administered only to Path A.
- P3 is administered to all three paths. Path B receives a stem explicitly asking about chosen rather than required reading; Paths A and C receive the general relationship-with-reading stem. The response scale is identical across paths.
- P3A (“People like me are readers”) is a separate diagnostic and is not scored in the TRHI.
- P12 is administered to Paths A and B only. Path C therefore has no observed P12 response in Wave 1.
- P13 family reading is a satellite measure and is not scored in the TRHI.
- P14A and P14 are administered to all paths, allowing civic perception and open-ended recommendations to include non-readers as well as readers.
- The instrument was administered in a fixed order: all public P-series items first, then the D-series demographics, then the nonpublic module last. Every scored item and the P14 open end was therefore collected before any respondent saw a demographic or nonpublic question. Question and response-option order was fixed rather than randomized in Wave 1. Longitudinal comparability therefore requires preserving that order unless a future methodology revision explicitly evaluates the effect of changing it.

TRHI Scoring Methodology

4.1 Component Architecture

The TRHI is a weighted composite of five respondent-level components. Each component is scored on a 0–100 rubric. Path-specific scoring rules are applied first; the population component score is then the weighted mean across all retained respondents. The population TRHI is the weighted mean of respondent-level composite scores, which is algebraically equivalent to applying the component weights to the weighted population component means.

COMPONENT	WEIGHT	SOURCE	FIELDDED SCOPE	POPULATION TREATMENT
Reading Frequency	35%	P1	Path A only	Observed P1 for A; structural zero for B/C
Reader Identity / Relationship	25%	P3	All paths	Observed P3 for A/B/C
Social Reading	15%	P11	Path A only	Observed P11 for A; structural zero for B/C
Ambient Cultural Engagement	10%	P12 items 1–4	Paths A+B	Observed P12 for A/B; structural zero for C
Local Ecosystem Participation	15%	P12 items 5–9	Paths A+B	Observed P12 for A/B; structural zero for C

Table 2. Component architecture and weights.

The five component weights remain 35/25/15/10/15. This is a design architecture rather than an empirically estimated causal model. The weights prioritize direct reading behavior and relationship with reading while retaining meaningful social, ambient, and ecosystem dimensions.

4.2 Component Scoring Rubrics

P1 — READING FREQUENCY (35%)

RESPONSE	SCORE
None	0
Less than 1 hour/week	20
1 to 2 hours/week	40
2 to 4 hours/week	60
4 to 7 hours/week	80
More than 7 hours/week	100

Table 3. Reading frequency scoring.

Interpretation caution: The P1 rungs are an ordinal scoring rubric. Any external literature used to contextualize a reading-time threshold does not validate an equivalent threshold for the composite TRHI.

P3 — READER IDENTITY / RELATIONSHIP (25%)

RESPONSE	SCORE
Reading isn't part of my life right now	0
I read occasionally when something catches my interest	33
Reading is a regular part of my life	67
Reading is something I prioritize and do frequently	100

Table 4. Reader identity scoring.

For longitudinal continuity, P3 remains the scored identity component. Technically, however, P3 contains both self-relationship and behavioral language. Public interpretation should therefore avoid presenting P3 as a validated pure self-identity scale. P3A remains a separate belonging/identity diagnostic.

P11 — SOCIAL READING (15%)

RESPONSE	SCORE
Never	0
1–2 times in the past 12 months	25
Several times in the past 12 months	50
Monthly or more	75
Weekly or more	100

Table 5. Social reading scoring.

P11 measures reading shared outward — recommending a book to someone or discussing one with others. It is distinct from the P12 ambient item on *seeking* recommendations, which measures reception rather than exchange. The rungs are frequency bands, not measures of depth or reach: the instrument records how often a respondent talked about a book, not with whom, in what setting, or to what effect.

P12 AMBIENT TIER — CULTURAL ENGAGEMENT (10%)

Four items are scored proportionally: 0/25/50/75/100 for 0/1/2/3/4 selected activities. Items are reading a book review/literary essay/literary news; seeking book recommendations; listening to a literary podcast; and following book-related social media accounts or communities.

P12 ECOSYSTEM TIER — LOCAL PARTICIPATION (15%)

Five items are scored proportionally: 0/20/40/60/80/100 for 0/1/2/3/4/5 selected activities. Items are attending a literary event/author talk; visiting a bookstore; visiting a library; attending a book festival; and participating in a community reading program. The P12 book-club item remains unscored in the ecosystem tier to avoid overlap with P11 social reading.

4.3 Path-Specific Structural Scoring

Wave 1 population scoring uses the following explicit rule set:

PATH	FREQUENCY	IDENTITY / RELATIONSHIP	SOCIAL	AMBIENT	ECOSYSTEM
A — Chosen Reader	Observed P1	Observed P3	Observed P11	Observed P12	Observed P12
B — Required-Only	0 structural floor	Observed P3	0 structural floor	Observed P12	Observed P12
C — Non-Reader	0 structural floor	Observed P3	0 structural floor	0 structural floor	0 structural floor

Table 6. Path-specific structural scoring.

The Path C Ambient and Ecosystem zeros are structural assumptions, not observed responses: Path C was not administered P12 in Wave 1. The convention is retained for the canonical Wave 1 TRHI so that the population index represents the full adult population. Future waves may administer P12 to Path C as an unscored diagnostic while preserving the legacy structural-zero series for longitudinal comparability.

4.4 Path C Structural-Zero Sensitivity

The Path C Ambient and Ecosystem zeros are the least observationally grounded element of the Wave 1 scoring, so their influence on the headline is bounded here. Path C is 6.68% of the weighted population, and the two affected components carry 25% of the index weight. The canonical Wave 1 score is unchanged; the alternatives below are non-canonical sensitivity scenarios only.

TREATMENT OF PATH C AMBIENT / ECOSYSTEM	POPULATION TRHI	CHANGE
Structural zeros — canonical Wave 1	39.70	—
Assigned Path B observed means (22.01 / 12.02)	39.97	+0.27
Assigned population means (24.65 / 20.94)	40.08	+0.37
Components computed on observed cases only	40.10	+0.40

Table 7. Path C structural-zero sensitivity (non-canonical scenarios).

Path B is the nearest observed comparison group. The rows above are illustrative observed-comparator scenarios rather than bounds; they move the index by no

more than 0.40 points. A bound requiring no assumption about non-reader behavior is available directly: Path C is 6.68% of the weighted population and the two unobserved components carry 25% of the index weight, so even assigning every Path C respondent a perfect 100 on both could raise the index by at most approximately 1.67 points. Both figures sit well inside the interpretation guardrail that treats movement under three points as broadly stable. The structural-zero convention is retained for the canonical series, and this exhibit is published so the size of the assumption is visible rather than asserted.

4.5 Score Bands — Public Interpretive Ranges

Wave 1 uses five civic interpretive ranges: Developing (0–19), Emerging (20–44), Healthy (45–59), Strong (60–74), and Exceptional (75–100). Bands are assigned on half-open intervals — a score falls in Emerging if it is at least 20 and below 45 — so a score of 44.70 is Emerging. The ranges are intended to make the 0–100 composite legible to nontechnical audiences and to provide a stable civic vocabulary for longitudinal reporting. They are not externally validated clinical thresholds or nationally normed states of population reading health.

RANGE	LEGACY LABEL	PUBLICATION TREATMENT
0–19	Developing	Public interpretive range
20–44	Emerging	Public interpretive range
45–59	Healthy	Public interpretive range
60–74	Strong	Public interpretive range
75–100	Exceptional	Public interpretive range

Table 8. Score bands and publication treatment.

Wave 1 public reporting may display the numeric score and interpretive range together—for example, “39.70 — Emerging”—provided the report identifies the band as an interpretive range. The numeric TRHI remains the precise result and the longitudinal anchor; movement across a band boundary does not by itself establish statistically significant or substantively large change.

4.6 TRHI Formula

For respondent i :

$$TRHI_i = 0.35F_i + 0.25I_i + 0.15S_i + 0.10A_i + 0.15E_i$$

The population TRHI is the final survey-weighted mean of $TRHI_i$ across all 511 retained respondents. The Path A TRHI is the final survey-weighted mean among Path A respondents only.

4.7 Pre-Fielding Simulation — Historical Design Check

Before fielding, the project used Monte Carlo simulations based on external priors to stress-test the scale and component architecture. Those simulations were design exercises, not estimates of Pima County reading health and are not used as evidence for Wave 1 findings. This methodology therefore retains them only as historical design documentation and does not use simulated component gaps to define the public narrative.

4.8 Final Wave 1 Score Result

MEASURE	UNWEIGHTED	FINAL WEIGHTED
Population TRHI	39.69	39.70
Reading Frequency	45.75	46.13
Reader Identity / Relationship	53.63	54.15
Social Reading	30.14	29.41
Ambient Cultural Engagement	25.05	24.65
Local Ecosystem Participation	21.60	20.94
Path A TRHI	45.29	45.44

Table 9. Weighted and unweighted comparison.

The final weighting changes the population TRHI by 0.01 point. That stability is reassuring, but it does not eliminate the limitations of nonprobability recruitment or justify a conventional margin of sampling error.

Bilingual Instrument and Translation

5.1 Rationale

The Wave 1 instrument was fielded in English and Spanish. The bilingual design was intended to reduce language-related exclusion and improve accessibility to Pima County adults who would not be well served by an English-only survey.

ACS language-at-home statistics are useful demographic context, but they are not used as a Wave 1 weighting margin because the TSOR D4 wording (“primarily spoken in your home”) is not equivalent to the Census language-at-home construct and the non-English Wave 1 cells are thin.

5.2 Translation

Translator: Mrs. Camila Kohen, CT — ATA-certified English-to-Spanish translator and interpreter. The target dialect was Mexican Spanish appropriate to a Tucson-area adult audience. The translation process included native Mexican-Spanish review for naturalness and register.

5.3 Translation Decisions and Wave 1 Limitation

The saved source documents do not contain a complete item-by-item translation decision log. This methodology does not reconstruct one from memory. The methodology therefore documents the translator, dialect target, bilingual fielding architecture, and known limitation. Fieldwork was aligned to both English and Spanish and Spanish-preference panelists were included in the sample frame, as confirmed by the fieldwork vendor; four retained respondents completed the Wave 1 survey in Spanish. This means bilingual availability improved access but does not support standalone Spanish-language subgroup estimates in Wave 1.

Sample and Panel Methodology

6.1 Fielding Platform

Panel vendor: Cint Exchange (SOW Q-145663), with the survey programmed/administered through the Forsta/Decipher environment. Fielding ran from 30 July to 14 August 2026 and produced 520 qualified completes before project-level QC exclusions.

6.2 Recruitment and Sampling Classification

Cint Exchange is a digital research marketplace that connects surveys to respondents sourced through multiple integrated suppliers. The project record does not document a known, nonzero selection probability for every Pima County adult. Accordingly, the technical methodology classifies Wave 1 as a nonprobability online-panel sample. For general public communication, the shorter phrase “weighted online panel survey” is used unless the technical sampling classification itself is under discussion.

Respondent incentives were determined by the individual supplier panels from which respondents were sourced, rather than by the project or by the fieldwork vendor; the vendor was unable to provide the specific incentive type or value for this fielding. Supplier composition and recruitment dispositions were requested and were not provided.

This classification governs inference; it is not a quality label. It does not mean that the survey is a casual internet poll or that demographic calibration is meaningless. It means that post-stratification can reduce known demographic imbalance but cannot establish the design-based selection probabilities required for a conventional probability-sample margin of sampling error or guarantee removal of selection bias on unmeasured characteristics.

6.3 Geographic Recruitment and Verification

Respondents were targeted to Pima County and supplied a home ZIP code (D1). Post-fielding verification found that the original operational target list omitted

legitimate Pima County ZIP codes 85744 and 85751. Respondents in confirmed Pima County ZIPs were retained even when the operational list was incomplete. The Wave 2 target ZIP reference must be re-verified before fielding.

6.4 Final Sample

PATH	FINAL N	UNWEIGHTED SHARE
Path A — Chosen Reader	431	84.3%
Path B — Required-Only Reader	48	9.4%
Path C — Non-Reader	32	6.3%
Total	511	100%

Table 10. Final analytic sample by path.

Wave 1 fielded 520 qualified completes. Nine were excluded after project-level quality/geography review: two for convergent quality concerns and seven for confirmed out-of-area geography. The final public and private analytic environments use the same retained respondent universe of 511.

6.5 Quotas and Representativeness

Earlier planning documents referenced an aspirational larger Path A sample; no hard path quota was imposed in fielding. The final Wave 1 publication methodology does not claim that a path quota or demographic quota converted the sample into a probability-representative sample. Returning respondents cannot be identified across waves, so an annual series on this frame should be understood as repeated panel-marketplace cross-sections with unknown overlap rather than as guaranteed independent samples. Population calibration is handled through the final raking protocol in Section 8. Any future wave that uses hard quotas should document the exact targets and achieved counts as part of that wave's methodology.

6.6 Subgroup Reporting Policy

Because Wave 1 uses a nonprobability panel and has N=511, subgroup analysis is descriptive rather than conventional design-based inference. The standing public reporting rule is:

- $n \geq 100$ unweighted: eligible for routine annual descriptive reporting, subject to substantive relevance and weighting diagnostics.
- $n = 50\text{--}99$: directional only; use sparingly and label the unweighted base prominently.
- $n < 50$: generally suppress as a standalone annual estimate. Consider pooled multi-year analysis instead.
- ZIP/neighborhood estimates are not published at Wave 1 sample size.

6.7 Wave-to-Wave Consistency

The measurement invariants to preserve are the wording, routing, order and scoring of the Wave 1 items that comprise the TRHI and its designated recurring diagnostics, path logic, scoring rubrics, component weights, geographic unit, core weighting targets, QC protocol, and public/private firewall. New questions may be added in future waves provided they do not alter those items or the sequence in which respondents encounter them. Panel supplier composition may vary between waves even where the fieldwork vendor does not. Supplier blend and respondent-incentive documentation were requested for Wave 1 and were not provided; that remains an open disclosure item, and any material change between waves will be disclosed when the information is available. Cross-supplier deduplication was requested and answered: the vendor described a layered mechanism operating across its supplier network, and the project independently checked device, session, and panelist identifiers across all 520 delivered completes and found no duplicates (Section 7.5).

6.8 Wave-to-Wave Respondent Overlap

Wave-to-wave respondent overlap cannot be measured in this design. A returning-respondent indicator was requested during Wave 1 setup; the vendor confirmed that its platform can flag which wave a record belongs to but has no mechanism to detect whether the same respondent appears in more than one wave. Suppression of prior-wave respondents is available and was declined: at Pima County's panel scale, excluding Wave 1 participants from a later wave would reduce the available frame below what that wave requires. Overlap is therefore structurally uncontrolled, and is disclosed as a limitation rather than corrected. The absence of cross-wave suppression is not, by itself, grounds to abandon the annual design.

Data Quality Control and Analytic Governance

7.1 Quality Review Standard

All vendor-flagged records were reviewed rather than automatically excluded. A single fast completion, unusual open end, or high selection count was not sufficient. Exclusion required converging evidence from multiple independent signals.

- Content plausibility: whether an open-ended response can reasonably be connected to the question, including plausible misreading.
- Generative dwell time: time on the open-ended page relative to the respondent's own page-time pattern rather than a single survey-wide cutoff.
- Selection-pattern consistency: whether unusually high/low checkbox selection behavior persists across unrelated topics.

The standard actually applied in Wave 1 was two independent problems in the same record before content/quality exclusion.

7.2 Speeder Review

TSOR paths differ materially in length, so a single blended completion-time threshold is not appropriate. Project review used path-specific timing context, including one-third of the path median as a screening threshold, followed by respondent-level content/pattern review rather than automatic deletion.

7.3 Geographic Verification

Self-reported D1 ZIP was checked against actual Pima County geography rather than only the original operational ZIP list. Seven qualified completes were excluded for confirmed geography outside Pima County.

7.4 Wave 1 Exclusion Result

520 qualified completes → 9 exclusions → 511 retained. The two quality exclusions can be justified from public P-series content and permitted survey paradata alone. This matters because the composition of the public TSOR sample must not depend

on responses that are excluded from public analysis.

7.5 Duplicate/Fraud Check

The project checked available respondent/device/session identifiers across all 520 qualified completes and found no duplicates on the inspected fingerprint/session/RID fields. Vendor descriptions of Cint’s cross-supplier quality architecture are treated as corroborating context rather than a project-specific audit result. No names, email addresses, or other direct personal identifiers were received. The panel, session, and device identifiers used for these checks are held in the analytic record and are not released.

7.6 Raw-Data Immutability and Backcoding

The vendor-delivered 520-case source file is an immutable archive. No original response is overwritten. Any recoding occurs only in derived analytic fields with record ID, original value/text, analytic target, confidence, rationale, and version retained in an audit log.

- Only HIGH-confidence direct semantic matches are eligible for analytic backcoding to an existing coded response.
- Plausible but interpretive mappings remain Other.
- If an Other open end merely repeats an already-selected category, the derived residual-Other flag can be removed without double-counting the already-selected category.
- P14 is not backcoded into another closed question. It is coded under its own multi-label thematic codebook.
- In Wave 1, twelve backcoding decisions were applied across four questions; two of them changed published figures.

7.7 P14 Qualitative Coding

P14 (“What would make Tucson a stronger reading city?”) was administered to all 511 retained respondents. 479 returned non-blank text, where non-blank means at least one non-whitespace character; 32 returned none. The blind coding package comprised 481 records, two of which contained blank P14 text, were scored

NO_SUGGESTION by both passes, and are excluded from the coded universe, leaving 479 coded responses. Wave 1 uses frozen Codebook v1.0 with multi-label thematic coding, meaning one response can receive multiple substantive themes. Original response text remains the source of truth. All 479 responses were coded independently in two separate large language model passes using the same frozen codebook; the second pass was blind to the first pass's response-level assignments and theme frequencies. The complete multi-label code set matched exactly on 76.0% of responses; across all 8,622 response-by-theme yes/no decisions, agreement was 98.0%. Raw agreement across a sparse multi-label set is high almost mechanically, because most of the 18 possible labels are correctly “no” for any given response. Two stricter statistics are therefore reported alongside it: positive agreement on the labels actually assigned was 86.3%, and Cohen's kappa at the decision level was 0.852. All 115 responses with any coding disagreement were re-adjudicated against the written codebook: 112 were resolved by applying the codebook's boundary rules, and 3 genuinely ambiguous cases were escalated to and approved by the project principal. The 364 responses on which both passes agreed exactly were not individually re-reviewed. This process is described as two independent LLM coding passes with human adjudication of all disagreements, not as traditional human inter-rater reliability.

Reporting rules for P14. Percentages are computed on the final analytic weights, not on unweighted counts. Theme shares are reported against the responses containing a substantive suggestion. Because a response may carry more than one theme, theme shares do not sum to 100%. The residual categories — no suggestion, current assets already adequate, and off-topic — are reported against the broader non-blank universe of 479 responses. Any published open-end percentage states which of those two denominators it uses.

7.8 Public/Private Analytical Firewall

Public TSOR analysis is restricted to P-series responses, D-series demographics, allowed QC/paradata, public contextual data, and derived public fields such as TRHI scores and survey weights. T-series responses and any T-derived finding are physically excluded from the public analytic dataset. A separate nonpublic analysis may use P, D, T, and other appropriate sources. Both products use the same 511 retained respondent universe but different evidence allowlists.

Post-Stratification Weighting

8.1 Purpose and Boundaries

Raking aligns the sample with selected ACS demographic margins. It can reduce imbalance on measured calibration variables; it cannot remove selection bias associated with unmeasured characteristics or turn a nonprobability panel into a probability sample.

8.2 Final Calibration Variables

MARGIN	WAVE 1 CATEGORIES	UNIVERSE	ACS SOURCE
Age	18–24; 25–34; 35–44; 45–54; 55–64; 65–74; 75+	All retained adults	2024 ACS 5-Year B01001
Education	<HS; HS/GED; some college/associate; bachelor+	Age 25+ with substantive D6	2024 ACS 5-Year B15003
Hispanic origin	Hispanic; non-Hispanic	Substantive D3 Yes/No	Adult 18+ share derived from 2024 ACS B01001I

Table 11. Raking margins and ACS sources.

Home language is not a Wave 1 weighting margin. The survey and ACS language questions are not equivalent, and the Wave 1 non-English cells are too thin for stable multi-category calibration. Sex/gender is not a weighting margin because TSOR did not collect it.

8.3 Final Raking Procedure

1. Initialize every retained respondent with weight 1.0.
2. Rake sequentially to the seven-category age margin across all retained respondents.
3. Rake education among age-25+ respondents with substantive D6 to the four B15003 categories. Respondents age 18–24 and education Prefer-not-to-say remain in the sample and are calibrated through their other applicable margins.

4. Rake Hispanic origin among respondents with substantive D3 Yes/No to the adult 18+ Pima County target. Prefer-not-to-say remains in the sample and is calibrated through other applicable margins.
5. Cycle margins until the maximum absolute calibration residual is below 1×10^{-10} .
6. Normalize weights to mean 1.0, so final weights sum to 511.
7. Apply a prospective maximum normalized weight cap of 3.0 if needed. No Wave 1 case reaches the cap, so no trimming is applied.

8.4 Calibration Balance

The table below shows what the weighting corrected. The achieved column confirms convergence rather than serving as a diagnostic: raking ran to a maximum absolute residual below 1×10^{-10} , so achieved and target agree to more decimal places than are printed. The information is in the unweighted column.

MARGIN	CATEGORY	UNWEIGHTED	ACS TARGET	ACHIEVED
Age	18-24	10.76%	14.40%	14.40%
Age	25-34	19.18%	16.22%	16.22%
Age	35-44	19.18%	14.48%	14.48%
Age	45-54	17.81%	13.31%	13.31%
Age	55-64	13.70%	14.74%	14.74%
Age	65-74	13.11%	15.02%	15.02%
Age	75+	6.26%	11.84%	11.84%
Education 25+	Less than high school	4.88%	9.48%	9.48%
Education 25+	High school / GED	29.27%	21.12%	21.12%
Education 25+	Some college / associate	40.13%	32.61%	32.61%
Education 25+	Bachelor's or higher	25.72%	36.79%	36.79%
Hispanic origin	Hispanic	32.94%	32.69%	32.69%
Hispanic origin	Non-Hispanic	67.06%	67.31%	67.31%

Table 12. Calibration balance: unweighted sample, ACS target, and achieved weighted distribution.

The largest pre-weighting imbalances were adults with a bachelor's degree or higher (11.07 points under-represented) and adults 75 and older (5.58 points under-represented). Hispanic origin required almost no correction (0.25 points). Prefer-not-to-say responses were retained and calibrated through their other applicable margins: 7 of 511 on Hispanic origin (D3) and 7 of 511 on education (D6).

8.5 Final Weight Diagnostics

DIAGNOSTIC	WAVE 1 VALUE
Iterations to convergence	11
Mean weight	1.000
Minimum	0.505
Median	0.908
95th percentile	1.867
Maximum	2.419
Cases > 3.0	0
Kish effective sample size	425.5
Unequal-weight design effect	1.201

Table 13. Final weight diagnostics.

The weighted population TRHI is 39.70 versus 39.69 unweighted. Weighting therefore matters for correct demographic calibration and item estimates even though it has almost no effect on the headline composite in Wave 1.

8.6 Weighting Limitations

- Calibration targets address only selected observed demographics.
- Education calibration uses the ACS 25+ universe and therefore does not force respondents age 18–24 into a 25+ population distribution.
- D3 and D6 Prefer-not-to-say cases are retained and calibrated through other applicable margins rather than imputed solely to satisfy a benchmark.
- The effective sample size is below 511 because unequal weights add variance.
- No claim is made that calibration eliminates all panel self-selection or coverage differences.

Theoretical Framework and Civic Purpose

9.1 Independent Development and Convergent Academic Precedent

One prior publication proposes both the concept of “reading health” and a “Reading Health Index” under those names, and it is cited here so readers can check the distinction rather than discover it independently. The TRHI was developed for the Tucson/Pima County context, and its architecture was substantially complete before the project identified Gerčar and van der Weel (2023), “Reading Health and the Reading Health Index.” That paper is a concept paper: it proposes a multi-source societal index drawing on literacy, publishing, bookselling, library and socioeconomic data, and does not field one. The TRHI is a survey-derived behavioral index of five components, with external data held outside the composite as context. TSOR is not derived from that paper, does not implement its framework, and should not be described as doing so. No other index informed it.

9.2 Operational Construct Model

The published Wave 1 methodology relies on the operational constructs measured in the instrument: reading frequency, reader identity/relationship, social reading, ambient cultural engagement, and local ecosystem participation, with family reading and civic perception retained as diagnostics outside the scored composite.

The component scales are not psychometrically equated latent-trait measures. A 10-point difference between two different component scales should not be interpreted as the same quantity of underlying “health.” Cross-component profiles are descriptive under the TSOR scoring rubrics; longitudinal change within a component is the cleaner comparative use.

9.3 Civic Purpose Statement

TSOR is designed to make reading culture visible as a civic condition rather than to diagnose individual readers. The project’s normative civic vision is that reading should be accessible, socially supported, culturally visible, and available across demographic groups. That vision informs why the measured dimensions matter, but it is not itself an empirical finding from the survey.

Known Limitations and Disclosure Framework

10.1 Sample Limitations

- Wave 1 is treated as a nonprobability online-panel sample because project documentation does not establish known selection probabilities for the Pima County adult population.
- Weighting aligns selected demographics but cannot correct unobserved panel self-selection, internet/panel coverage differences, or all supplier-blend effects.
- The final N=511 supports countywide descriptive reporting but limits fine-grained demographic, geographic, and language subgroup analysis. Wave 1 also did not collect sex/gender, so no gender-based reading analysis or sex/gender calibration is possible in the baseline edition.
- Only four retained respondents completed the survey in Spanish; Spanish-language subgroup statistics are not published. Hispanic-origin weighting does not correct undercoverage of Spanish-dominant adults: origin and language are distinct variables, and the Wave 1 frame carries no language calibration.
- The original operational ZIP target list omitted legitimate Pima County ZIPs, requiring post-fielding geographic verification. This was corrected analytically but is an operational limitation to fix before Wave 2.
- The weighted effective sample size is approximately 425.5 because of unequal weighting.

10.2 Instrument and Index Limitations

- P0 measures broad voluntary reading/listening, including books, articles, audiobooks, scripture, and other material. It should not be reported as “book reading” prevalence.

- P5, which asks how respondents typically obtain the reading material they use most, phrases three of its five response options in purchase terms (“browse and buy,” “go somewhere in person to buy,” “usually buy online”). Respondents whose primary route is library borrowing have no exactly matching option and may select the mixed-pattern response or the nearest in-person option. The P5 distribution, and the travel-time subgroup routed from it, should be read with that in mind. P5 is not a scored TRHI item and no component or index value is affected. The option set is a Wave 2 instrument correction.
- P3 blends reading relationship/identity with behavioral wording and is not a validated pure self-identity scale. P3 also uses path-specific stem wording: Path B receives a stem referring explicitly to chosen rather than required reading, while Paths A and C receive the general relationship-with-reading stem. The response scale is constant across all paths, but the item is not textually identical for every respondent, and the component carries 25% of the index.
- Path C was not administered P12. Its Ambient and Ecosystem structural zeros are a scoring convention, not observed nonparticipation.
- The five component scales use different item structures and difficulty. Cross-component point differences should not be interpreted as psychometrically equivalent distances.
- The qualitative score bands are interpretive civic ranges rather than externally validated thresholds.
- All survey behavior is self-reported and therefore subject to recall, interpretation, and social-desirability effects.
- Question/response order was fixed, so order effects cannot be separately estimated in Wave 1.

10.3 Statistical Uncertainty, MOE, and Precision Orientation

Wave 1 does not report a conventional design-based margin of sampling error as though the 511 cases were drawn by simple random probability sampling. The design does not establish the selection probabilities required for that claim. However, because readers and journalists commonly ask for an MOE-scale reference, TSOR may provide a clearly labeled sampling-only orientation benchmark: a simple random probability sample of N=511 would have a maximum

95% margin of sampling error of approximately ± 4.3 percentage points — “maximum” because sampling error is widest for a result near 50% and narrower for results near 0% or 100%; applying Wave 1’s unequal-weight design effect of 1.201 yields an analogous sampling-only benchmark of approximately ± 4.8 points. These values are not the formal MOE of the online-panel design and do not quantify panel self-selection, coverage, measurement, or other nonsampling error. A description of TSOR as carrying a margin of error of about five points is a misreading of this paragraph.

The standard public answer when an MOE is requested is therefore: “A conventional margin of error does not strictly apply to this weighted online-panel design. For scale, a random sample of this size would be roughly ± 4.3 percentage points at 95% confidence, or about ± 4.8 points after accounting for the Wave 1 weighting design effect. That benchmark applies to a single percentage near 50%. The TRHI is a weighted composite of five component scores and varies less than any one proportion; the same calculation applied to the composite gives roughly ± 2 points. Neither is a margin of error for this study.” Exact unweighted subgroup bases, weight diagnostics, sensitivity analysis, and longitudinal guardrails should accompany this orientation rather than converting it into a false probability-sample claim.

10.4 External Comparability

TSOR has no direct national city-level equivalent. External sources such as Pew, NEA, ATUS, PIAAC, ACS, PCPL/IMLS, and Arizona library statistics measure different constructs, periods, universes, or geographies. Public reporting should label each comparison as Benchmark, Context, or Not directly comparable and should never fold an external statistic into the TRHI after the fact.

10.5 Standard Publication Disclosure — Wave 1

Recommended disclosure: “Wave 1 of the Tucson State of Reading was fielded from 30 July to 14 August 2026 through Cint Exchange as a weighted online panel survey of Pima County adults age 18 and older. Of 520 qualified completes, 9 were excluded after documented quality and geographic review, yielding an analytic sample of 511. Results are post-stratification weighted to 2024 ACS Pima County benchmarks for age, educational attainment, and Hispanic origin. Because respondents were

recruited through an online panel rather than selected from a probability sample of all county adults, a conventional margin of sampling error is not reported. For orientation only, a random sample of this size would have a maximum 95% sampling margin of approximately ± 4 –5 percentage points. Weighting can reduce known demographic imbalance but cannot eliminate selection bias on unmeasured characteristics. Public TSOR findings use only the public P-series, D-series demographics, permitted QC/paradata, and derived public fields; private T-series responses are excluded from public analysis.”

10.6 Annual Change Interpretation Guardrail

Before Wave 2 results are examined, TSOR adopts a simple longitudinal interpretation guardrail for the headline TRHI: changes under 3 points are described as broadly stable unless multiple independent indicators show a coherent pattern; changes of 3–5 points are described as notable movement requiring component-level review; changes greater than 5 points are described as substantial movement subject to the same methodological review. These are editorial and longitudinal interpretation guardrails—not statistical-significance thresholds, confidence intervals, or a substitute for examining component and item-level evidence.

SECTION ELEVEN

Methodology Decisions — Formal Changelog

Guidance on appropriate and inappropriate use of TSOR findings is published separately, in the *Findings and Reporting Guide*, available from The Place & Page. It is not reproduced here.

Status values used below. LOCKED — fixed for this and all future waves. LOCKED W1 — fixed for the Wave 1 record; may be revisited for later waves except where Section 6.7 lists the item as a measurement invariant. CORRECTED — a Wave 1 decision that revised an earlier understanding. PUBLICATION RULE — governs how a result may be presented publicly. GOVERNANCE — a standing rule for how the project operates across waves.

DATE	DECISION	STATUS
June 2026	Five-component TRHI architecture retained: Frequency, Reader Identity/Relationship, Social, Ambient, Ecosystem.	LOCKED
June 2026	Component weights 35/25/15/10/15.	LOCKED
June 2026	P13 family reading remains satellite, not scored.	LOCKED
June 2026	Native 0–100 score reporting; baseline=100 index rejected.	LOCKED
Aug. 17, 2026	Final analytic universe fixed at N=511 after 9 documented exclusions.	LOCKED
Aug. 17, 2026	P3 scope corrected to all paths based on final fielded programming specification.	CORRECTED
Aug. 17, 2026	Path C P12 treatment documented as structural zero, not observed response.	CORRECTED
Aug. 17, 2026	Final raking protocol: age + education (25+) + adult Hispanic origin; no home-language or sex calibration.	LOCKED W1
Aug. 17, 2026	Final weighted population TRHI = 39.70; Path A = 45.44.	LOCKED W1

Aug. 17, 2026	Five score bands retained for public use as interpretive civic ranges; numeric TRHI remains the precise longitudinal anchor.	PUBLICATION RULE
Aug. 17, 2026	No formal conventional MOE is claimed; public materials may provide a clearly labeled random-sample-equivalent orientation of approximately ± 4 –5 percentage points.	PUBLICATION RULE
Aug. 17, 2026	Raw source archive immutable; only high-confidence derived backcoding allowed.	LOCKED
Aug. 17, 2026	P14 Codebook v1.0 established as multi-label qualitative baseline.	LOCKED W1
Aug. 17, 2026	Public TSOR/T-series firewall physically implemented.	LOCKED
Aug. 17, 2026	P3 retained as the scored Reader Identity / Relationship component; P3A retained as a separate Reader Belonging / Social Identity diagnostic; no P3/P3A averaging.	LOCKED W1
Aug. 17, 2026	P14 Codebook v1.0 finalized through two independent LLM coding passes; all 115 disagreements adjudicated; three ambiguous cases approved by project principal.	LOCKED W1
Sept. 2026	Weighting-vintage rule fixed for future waves: same calibration dimensions, universes, and categories as Wave 1, with benchmark values updated to the then-current ACS 5-year release.	GOVERNANCE
Aug. 17, 2026	Annual TRHI interpretation guardrail prospectively fixed before Wave 2: <3 broadly stable; 3–5 notable; >5 substantial, subject to multi-indicator review; not a significance test.	GOVERNANCE

Table 14. Methodology decision record.

APPENDIX A

Final Weighting Benchmarks and Diagnostics

A.1 Age — Pima County Adults 18+

AGE GROUP	ACS COUNT	TARGET SHARE
18–24	122,759	14.40%
25–34	138,360	16.22%
35–44	123,440	14.48%
45–54	113,505	13.31%
55–64	125,712	14.74%
65–74	128,045	15.02%
75+	100,947	11.84%

Table A1. Age benchmarks.

A.2 Education — Pima County Population 25+

EDUCATION	ACS COUNT	TARGET SHARE
Less than high school	69,181	9.48%
High school/GED	154,207	21.12%
Some college / associate	238,020	32.61%
Bachelor's degree or higher	268,601	36.79%

Table A2. Education benchmarks.

A.3 Hispanic Origin — Adults 18+

CATEGORY	ACS COUNT	TARGET SHARE
Hispanic / Latino	278,748	32.69%
Non-Hispanic	574,020	67.31%

Table A3. Hispanic origin benchmarks.

A.4 Final Weight Diagnostics

METRIC	VALUE
Mean weight	1.000
Minimum	0.505
Median	0.908
95th percentile	1.867
Maximum	2.419
Kish effective N	425.5
Unequal-weight design effect	1.201

Table A4. Weight diagnostics.

APPENDIX B

Final Wave 1 TRHI Score Summary

MEASURE	WEIGHTED SCORE
Population TRHI	39.70
Reading Frequency	46.13
Reader Identity / Relationship	54.15
Social Reading	29.41
Ambient Cultural Engagement	24.65
Local Ecosystem Participation	20.94
Path A TRHI	45.44

Table B1. Wave 1 TRHI score summary.

Public/Private Data Firewall and Open-End Coding Governance

C.1 Public TSOR Allowlist

- Public P-series survey responses and derived public question metrics.
- D-series demographics and final Wave 1 survey weights.
- Permitted QC and paradata needed to justify inclusion/exclusion and data integrity.
- TRHI respondent-level and aggregate scores derived solely from public inputs.
- Public contextual sources (Census, library statistics, NEA, Pew, ATUS, PIAAC, and similar) when clearly sourced and comparability-labeled.

C.2 Public TSOR Exclusions

- All T-series survey responses.
- Any theme, segment, recommendation, or editorial emphasis derived from T-series response content.
- Any exclusion decision that cannot independently be justified from public responses/permitted QC evidence.

C.3 Open-End Quotation Governance

The fielded survey preamble, displayed to every respondent before the first question and immediately above the consent control, states: “Your responses are completely anonymous. No personally identifying information will be collected or shared. Results will be reported only in the aggregate.”

That commitment is a constraint on the form of reporting, not only on identifiability. **Verbatim open-end responses are therefore not reproduced in any public TSOR output** — not in the annual report, not in partner or media materials, and not in derived products — regardless of whether an individual response could be identified. Open-end material is published only as coded theme frequencies and as aggregate characterization.

Original response text is retained in the analytic record for coding, adjudication, and audit. It is not released, quoted, or excerpted.

Any future wave that intends to publish verbatim responses must first obtain that permission in the fielded preamble of the wave in which those responses are collected. Consent given under the Wave 1 preamble cannot be extended retroactively.

C.4 Backcoding Audit Fields

Every derived recode should preserve: record identifier, original coded choice(s), original open text, proposed analytic category, confidence, action, rationale, coder/version, and review status. Raw data are never overwritten.

References and Source Notes

- Pima County Reading Culture Survey — Survey Programming Specification — Cint / Forsta, July 2026. Authoritative source for fielded routing and response options.
- TSOR Wave 1 Public Evidence FINAL v1.4 workbook, August 17, 2026. Canonical retained analytic sample, final respondent-level weights, TRHI scoring, QC/backcoding logs, two-pass P14 coding and adjudication, final theme frequencies, public/private firewall, and report blueprint.
- U.S. Census Bureau, 2024 ACS 5-Year Estimates: B01001 (Sex by Age), B15003 (Educational Attainment for Population 25+), B01001I (Sex by Age, Hispanic or Latino).
- AAPOR, Transparency Initiative / guidance on probability and nonprobability sample disclosure and sampling-error reporting. <https://aapor.org/standards-and-ethics/transparency-initiative/>
- AAPOR, Data Quality Metrics for Online Samples: Considerations for Study Design and Analysis (2023). <https://aapor.org/reports/data-quality-metrics-for-online-samples-considerations-for-study-design-analysis/>
- Cint, Cint Exchange and respondent sourcing documentation. <https://www.cint.com/products/cint-exchange/> and <https://www.cint.com/solutions/source-respondents/>
- Gerčar, J., & van der Weel, A. (2023). “Reading Health and the Reading Health Index.” *Societies*, 13(4), 86. <https://doi.org/10.3390/soc13040086> — TSOR is not derived from this framework and does not claim derivation from it.

Disclosure tiers: The public TSOR report carries the scoring rubrics, index formula, component weights, weighting targets, sample disclosure, and limitations. This companion volume records the fielded instrument architecture, quality-control protocol, weighting mechanics, and dated decision record, and is published alongside the report at theplaceandpage.com/tsor/2026. The fielded questionnaire wording and response options, the survey programming specification, the P14 codebook and coding instructions, the scoring implementation, and audit logs are retained by The Place & Page. Respondent-level data, custom tabulations, and

verbatim open-ended responses are not released.

Wave 1.0 — September 2026. Wave numbering is *N.M*: N is the survey wave, M the document revision within it. Corrections after publication are documented in a public errata and version log maintained by The Place & Page.

© 2026 The Place & Page. Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0).

Recommended citation. The Place & Page. (2026). *Tucson State of Reading Methodology Documentation, Wave 1*. doi.org/10.5281/zenodo.22714565

Edition page: theplaceandpage.com/tsor/2026

Correspondence and corrections: tsor@theplaceandpage.com.

Version-control note: This document is the frozen Wave 1 publication methodology. Any future substantive change to the scoring architecture, component weights, Wave 1 weights, public/private firewall, final analytic N, or treatment of baseline results must be issued as a new version or explicit erratum rather than silently replacing this record.

TUCSON STATE OF READING

Methodology Documentation · Wave 1.0

Produced by The Place & Page · Tucson, Arizona · September 2026

Published alongside the report at theplaceandpage.com/tsor/2026