# Contents

# A Supplementary Material

This supplementary material documents the LLM-assisted coding pipeline, the codebook used to classify invocations, the prompts and output schema, the multi-layer validation protocol, the corpora and speaker metadata, robustness checks, full logistic-regression specifications, the Spanish-language originals of the block quotes that appear translated in the main text, and replication materials.

Theoretical framing was deliberately removed from every artifact the classifier saw. The system prompt, the codebook v2.7 sent as the cached system context, and the country-specific guidance contain no reference to "cherry-picking," "representative claim-making," or related interpretive vocabulary. The instrument is limited to procedural description, definitional criteria, segmentation rules, priority rules, and the decision procedure. This was a v2.5 revision motivated by concern that interpretive priming would compromise the textual diagnosability of Descriptive Reference as the residual category (see §A.2.10).

## A.1 Pipeline summary and coding passes

**Model and parameters.** The pipeline uses Anthropic's `claude-opus-4-6` as the classifier with temperature 0, `max\_tokens` 8192, and prompt caching on the codebook block (`cache\_control:\{"type":"ephemeral"\}`). Caching makes the marginal input cost of each speaker turn the country-specific guidance and the turn text only; the codebook—about 5,600 tokens of cached prefix—is paid once per cache window. Each turn receives an independent API call; classifications are not conditioned on neighboring turns.

**Coding passes.** Across codebook development and final analysis, the pipeline was executed across roughly 14 distinct passes summarized in Table 1. Pilots on earlier codebook versions (v2.2 on Opus 4.5; v2.5 and v2.6 on Opus 4.6) were run on corpus subsets to manage cost and to surface failure modes (see §A.2.10 for the changelog). The canonical v2.7 prompt was run once on the full analytic corpus to produce the predictions analyzed in the main text. Three additional v2.7 canonical runs and one paraphrased-prompt run on a 13-session Chile subsample plus the full

Table 1: Coding passes used in codebook development and final analysis.

| Pass | Codebook | Model | Variant | Corpus |
|---|---|---|---|---|
| Pilot v2.2 | v2.2 | Opus 4.5 | canonical | subsets |
| Pilot v2.5 | v2.5 | Opus 4.6 | canonical | subsets |
| Pilot v2.6 | v2.6 | Opus 4.6 | canonical | subsets |
| Run 1 (analytic) | v2.7 | Opus 4.6 | canonical | full Chile (80 sessions) + full Cuba (3 sessions) |
| Run 2 | v2.7 | Opus 4.6 | canonical | 13-session Chile subsample + full Cuba |
| Run 3 | v2.7 | Opus 4.6 | canonical | 13-session Chile subsample + full Cuba |
| Run 4 | v2.7 | Opus 4.6 | paraphrased | 13-session Chile subsample + full Cuba |

Cuba corpus feed the inter-run and inter-prompt reliability analyses in §A.4.4. Across all passes, approximately 56,700 speaker turns were coded; total API cost was just under $450.

**Per-row processing.** For each row of an input Excel file, the pipeline (`scripts/pipeline/code\_transcripts.py`) builds a user message with the row's `session\_id`, `row\_id`, country, and verbatim text; calls the API with the cached system prompt; strips any markdown code fences from the response; parses JSON; normalizes casing on `valence` and `mechanism`; and validates the output against the schema in §A.3.4. Schema-invalid responses are excluded from the analytic CSV and logged as `error` in `processing\_log.csv`. The 13-session subsample for cross-run comparison oversampled invocation-rich sessions (41% of Chile's positive turns) so classification-level $\alpha$ has adequate denominators.

**Run identifiers and provenance.** Every output row carries `run\_id` (timestamp + variant), `codebook\_version`, `prompt\_variant` (`canonical` or `paraphrased`), and `temperature`. The flat CSV `uid` key is `\{session\_id\}\_r\{row\_id\}\_i\{invocation\_id\}`, ensuring every classification is traceable from the analytic dataset back to the source turn and the run that produced it.

## A.2 Codebook (v2.7)

Coding Invocations of Public Input in Constitutional Deliberations

*LLM-Assisted Coding Protocol*

*Version 2.7 — April 2026*

### A.2.1 Overview

This codebook provides instructions for identifying and coding references to formal public consultation mechanisms in constitutional plenary debates. The coding task has two stages: (1) identifying discrete invocations of public input within speaker turns, and (2) classifying each invocation by rhetorical function and valence.

### A.2.2 Data Structure

Data consists of Excel files containing plenary session transcripts. Each row represents one speaker turn. Key columns include speaker identification variables, the full text of the speaker's intervention, and metadata (party, ideology scores, commission assignments, etc.). The LLM will process each speaker turn individually, identifying any invocations of formal consultation mechanisms within the text. Each row represents one speaker turn, which is the unit of observation for classification. The unit of analysis is the invocation (see §A.2.4): a single speaker turn may contain zero, one, or multiple invocations.

### A.2.3 What Counts as Formal Consultation

Code ONLY references to formal consultation mechanisms established by institutional procedures. This includes:

**Chile.**

**Iniciativas Populares de Norma (IPNs).**

Citizen-proposed constitutional provisions submitted through the official platform. References may include "iniciativa popular," "iniciativa ciudadana," specific IPN numbers, signature counts, or named initiatives.

**Audiencias públicas.**

Formal public hearings organized by thematic commissions. References include "audiencia pública," testimony from invited experts or civil society representatives, and commission hearing records.

**Consulta indígena.**

The special indigenous consultation process conducted under ILO Convention 169 standards. References include "consulta indígena," results from indigenous participation mechanisms, and proposals from indigenous peoples submitted through the formal process.

**Cabildos autoconvocados.**

Self-convened local assemblies whose results were systematized and submitted to the Convention. Code ONLY when referenced as formal input (e.g., "los resultados de los cabildos"), not as general popular mobilization.

**Encuentros/semanas territoriales.**

Official territorial outreach events organized by the Convention.

**Secretaría de Participación Popular.**

Official communications, systematized results, or reports from the Secretariat of Popular Participation.

**Official communications from organized groups.**

Formal submissions from civil society organizations, unions, professional associations, etc., channeled through official Convention procedures.

**Cuba.**

**Consulta popular.**

The mass consultation process for the 2019 Constitution. References include "consulta popular," "proceso de consulta," aggregate statistics from the consultation, and specific proposals collected.

**Propuestas ciudadanas.**

Citizen proposals submitted through the formal consultation process. May be referenced as "propuestas," "planteamientos," or "opiniones" when clearly tied to the official mechanism.

**Exclusions.** Do NOT code the following as formal consultation references:

**General mandate claims.**

References to "el pueblo," "la ciudadanía," "quienes nos eligieron," or abstract popular will without connection to specific consultation mechanisms.

**The estallido social.**

References to "octubre de 2019," "el estallido," "la revuelta," or the social uprising as legitimation source. These invoke popular mandate, not formal consultation.

**Electoral mandate.**

References to the speaker's own election, their district's preferences, or campaign promises.

**Media or public opinion.**

References to polls, media coverage, social media sentiment, or general public reaction unless explicitly tied to formal consultation results.

**Personal communications.**

Anecdotal references to conversations with constituents, personal testimony, or individual citizens unless framed as part of formal consultation.

**Historical consultation processes.**

References to past consultation processes (e.g., the 2016 cabildos under Bachelet, historical indigenous consultation under previous governments) that are not part of the current constitutional process. Code only references to the formal mechanisms established for the current deliberation.

### A.2.4 Unit of Analysis and Segmentation

The unit of analysis is a discrete invocation of formal public consultation. A single speaker turn may contain zero, one, or multiple invocations.

**Segmentation rules.**

**One invocation per distinct referent.**

If a speaker cites three different IPNs, code three separate invocations. If a speaker references both the consulta indígena and an audiencia pública, code two invocations.

**Continuous reference = one invocation.**

If a speaker discusses the same consultation mechanism across multiple sentences without interruption, code as one invocation. The rhetorical function applies to the entire continuous reference.

**Return after intervening material = new invocation.**

If a speaker references an IPN, discusses topically distinct material (see §A.2.4), then returns to the same IPN, code as two separate invocations.

**Generic + specific = code specific only.**

If a speaker makes a generic reference to "participación ciudadana" and then specifies "las iniciativas populares," code only the specific mechanism unless the generic reference serves a distinct rhetorical function.

**Segmentation by rhetorical move, not by sentence.** Segment by distinct rhetorical move within speaker turns, not by sentence. A speaker who cites multiple statistics about the same consultation topic in service of one rhetorical function (e.g., reporting the number of opinions, the percentage of total, and the number of meetings where it was raised) is making one invocation, not three.

*Segmentation triggers (new invocation):*

- *New substantive referent:* The speaker shifts to a different topic or issue from the consultation (e.g., from proposals on mandatory work to proposals on parental responsibility).

- *New rhetorical function:* The speaker shifts from one function to another with respect to the same or different input (e.g., from reporting statistics to explaining why the commission rejected the proposals).

*Not segmentation triggers (continuation of same invocation):*

- Additional sentences elaborating the same point.
- Additional statistics about the same issue.
- Rhetorical emphasis or repetition.

*Guiding question.* "Is this a new claim about a different piece of consultation input, or a new move with respect to the same input?" If yes, code as a new invocation. If no, treat as continuation of the current invocation.

*Unsegmentable cases.* When multiple rhetorical functions are blended within an utterance that cannot be cleanly segmented at a natural break, code the dominant or terminal function and include the complete utterance in the `text\_excerpt`. Do not artificially segment mid-sentence or produce fragmentary excerpts. The `text\_excerpt` should always be a complete, interpretable rhetorical unit—the full expression of the coded move from where the speaker introduces the referent to where they conclude the point.

*Sustained multi-dimensional evaluation.* When a speaker evaluates the same consultation mechanism across multiple sentences, touching on different dimensions of process quality (e.g., transparency, inclusiveness, participation rates, democratic value, political courage), this is one invocation. Different dimensions of the same evaluative move are elaboration, not new referents. The same applies to sustained substantive deployment—a speaker citing multiple statistics or examples in support of the same provision is making one invocation, not one per statistic.

*Priority of intervening material rule.* When a speaker returns to evaluate or deploy the same consultation mechanism after intervening material, apply the "return after intervening material" rule. Sustained multi-dimensional evaluation applies only to continuous passages without topical

interruption. See §A.2.4 for the definition of intervening material and §A.2.4 for the operational test.

**Defining intervening material.** The priority rule above depends on what counts as intervening material. Intervening material means *topically distinct* material—specifically, material that addresses a different referent than the one the speaker is using consultation input to address. The referent is the constitutional provision, article, procedural decision, or issue at stake in the speaker's current argument.

Mere elaboration, defense, extension, or restatement of the same argument does *not* constitute intervening material, even when it spans multiple sentences and does not itself reference consultation. A speaker building one sustained argument for or against a provision is making one rhetorical move, and consultation references threaded through that argument belong to one invocation.

**Operational test.** To determine whether a passage between two consultation references is intervening material or elaboration, apply the following test:

> *Could the passage between the two references be removed without changing what the speaker is arguing about?*

If **yes**, the passage is elaboration. The two consultation references belong to one invocation. Code the dominant or terminal rhetorical function and include the full span in the `text\_excerpt`.

If **no**—because the passage is itself about a different referent (a different article, a different provision, a different policy domain, a pivot to general legitimacy claims, or a walkthrough to the next item on the agenda)—the passage is intervening material. The return to consultation is a new invocation.

The test applies to the *intervening passage*, not to the consultation references themselves. This placement of the judgment is deliberate: asking "is this passage about the same thing?" is more tractable than asking "is this a new rhetorical move?"

**Worked examples.** *Example—one invocation (elaboration).* A speaker arguing for deferring the definition of marriage to the Código de Familia says *"el pueblo planteó que dejáramos esto para el código de familia,"* then develops the argument across several sentences about minority rights, the courage of submitting the question to popular decision, and the need to avoid the topic becoming central to the constitutional project, then returns with *"una vez que el pueblo se sintió con ese derecho adquirido, el 24% de las opiniones..."* The intervening passage elaborates the same deferral argument; removing it would not change what the speaker is arguing about. → **One invocation (Substantive Deployment / Positive).**

*Example—two invocations (topical interruption).* A speaker walks through the Cuba draft article by article, saying about article 84 *"esto es algo que se pidió mucho en la consulta popular,"* then moves to articles 85 and 86 on different topics (family violence and children's rights), then returns to an earlier article with *"y recuerden que sobre ese artículo también hubo muchas opiniones."* The passage between is about different articles; removing it would change the referent. → **Two invocations.**

### A.2.5 Coding Scheme

The coding scheme has two dimensions: Rhetorical Function (three mutually exclusive categories) and Valence (coded where applicable). For each identified invocation, assign exactly one Rhetorical Function and, if applicable, one Valence.

**Category 1: Descriptive Reference.**

**Definition.**

> Factual reporting of consultation activities, results, or procedures without evaluation, argumentation, or deployment for any rhetorical purpose. The speaker merely notes that consultation occurred or reports its contents neutrally.

**Valence.**

> Not applicable.

**Example (synthetic).**

*"La Secretaría de Participación ha informado que se recibieron 2,847 iniciativas populares de norma durante el período de inscripción."* [Pure factual report of consultation statistics without evaluation or deployment.]

**Category 2: Process Evaluation.**

**Definition.**

Assessing the quality, legitimacy, representativeness, or effectiveness of consultation mechanisms themselves. The speaker evaluates how consultation was conducted, not what it produced substantively. The evaluation of process is the terminal rhetorical move, not instrumental to another purpose.

**Key markers.**

Comments on timing, accessibility, methodology, representativeness, fairness, transparency, or institutional handling of consultation. Claims that consultation was adequate/inadequate, inclusive/exclusive, legitimate/illegitimate.

**Valence options.**

Positive (praising the process), Negative (criticizing the process), Mixed (acknowledging both strengths and weaknesses).

**Example—Positive (synthetic).**

*"Quiero agradecer el trabajo de la Secretaría de Participación y Consulta Indígena. Pese a los plazos acotados, lograron realizar un proceso de consulta que permitió la participación de miles de hermanos y hermanas indígenas."* [Evaluates the consultation process positively; does not deploy content to support a provision.]

**Example—Negative (synthetic).**

*"La consulta indígena se realizó en plazos extraordinariamente breves y con recursos insuficientes. No podemos pretender que este proceso cumpla con los estándares del Convenio*

*169."* [Criticizes the process; does not use this critique instrumentally to oppose a specific provision.]

**Category 3: Substantive Deployment.**

**Definition.**

Using the content of public input to support, oppose, or warrant specific constitutional provisions, procedural decisions, or topics. The speaker mobilizes what citizens said, proposed, or demanded through consultation mechanisms as an argument for or against a concrete normative position, to elevate the importance of an issue, or to justify procedural choices such as deferral.

**Key markers.**

Speaker cites consultation content (proposals, demands, testimony) in the context of arguing for or against an article, provision, or amendment, or uses consultation references to argue that an issue should be addressed, has been overlooked, or deserves priority. The consultation reference functions as evidence, legitimation, or warrant for a substantive position, procedural decision, or for elevating an issue's importance.

**Valence anchor.**

Valence for Substantive Deployment is coded with respect to the speaker's orientation toward the provision, decision, or issue being argued—not with respect to the speaker's orientation toward the input itself. A speaker may embrace the cited input while arguing against a provision (e.g., "you are ignoring 60,000 signatures"): this is Negative because the rhetorical move opposes the provision. A speaker may cite input to support a provision, elevate an issue for consideration, or warrant a procedural decision such as deferral: these are Positive.

**Valence options.**

Positive (deploying consultation content to support a provision, elevate an issue, validate a decision, or argue that a topic deserves consideration), Negative (deploying consultation

content to oppose a provision, warn against a direction, dismiss proposals as illegitimate, or argue that opponents have ignored public input), Mixed (see rule below).

*When to use Mixed.* Use Mixed only when both orientations are load-bearing in the same rhetorical unit—e.g., the speaker simultaneously supports one aspect of a provision and opposes another using the same consultation reference, or explicitly acknowledges tension in the input itself as part of the argument. If one orientation is dominant and the other is incidental or merely acknowledged in passing, code the dominant orientation rather than Mixed. Mixed is not a hedge for uncertainty; uncertainty is recorded in the confidence field.

*Example—Positive (synthetic). "El artículo 17 sobre derecho a las tierras indígenas recoge directamente las propuestas de la consulta indígena. Los pueblos originarios nos pidieron el reconocimiento de sus territorios ancestrales, y eso es exactamente lo que estamos votando hoy."* [Deploys consultation content to support the provision under debate.]

*Example—Positive/agenda-raising (synthetic). "Múltiples iniciativas populares y testimonios en audiencias públicas han planteado la necesidad de reconocer constitucionalmente los derechos de los animales. Este es un tema que la ciudadanía nos está pidiendo que abordemos."* [Deploys consultation content to elevate an issue for consideration. Coded as Positive because the speaker uses public input to argue in favor of addressing the topic.]

*Example—Positive/deferral (synthetic). "Nuestro pueblo planteó que dejáramos la definición del matrimonio para otro momento. Por eso la decisión fue llevarlo al Código de Familia, que se someterá a referéndum."* [Deploys consultation content to warrant a procedural decision (deferral). Coded as Positive because the speaker embraces the input approvingly.]

*Example—Negative (synthetic). "La iniciativa popular más firmada sobre este tema pedía exactamente lo contrario de lo que propone este artículo. Ustedes están ignorando la voz de más de 60.000 ciudadanos que se manifestaron a través del mecanismo oficial."* [Deploys consultation content to oppose the provision and criticize other drafters.]

*Example—Negative/dismissal (synthetic). "Hubo propuestas para eliminar el mercado de la Constitución, pero esas opiniones vienen de desconocedores de la realidad económica. No pode-*

*mos tomar en serio planteamientos que ignoran décadas de experiencia."* [Deploys consultation content to dismiss proposals. Coded as Negative because the speaker treats the input dismissively.]

**Priority rules.** When an invocation could fit multiple categories, apply the following rules:

**Rule 1 (Process → Substantive).**

If process evaluation is instrumental to substantive deployment—that is, the speaker evaluates the consultation process in order to support or oppose a provision—code as Substantive Deployment. Code as Process Evaluation only when assessment of consultation quality is the terminal rhetorical move.

**Rule 2 (Descriptive as residual).**

Code as Descriptive Reference only when no evaluative or argumentative function is present. If there is any discernible rhetorical purpose, code the appropriate function.

*Applying Rule 1: the premise test.* Rule 1 requires distinguishing process evaluation that is instrumental to substantive deployment from process evaluation that is a terminal rhetorical move. In sustained speaker turns, process praise or criticism and substantive arguments often coexist, and a mechanical test is needed to avoid arbitrary assignment.

Code as Substantive Deployment (per Rule 1) only when the process claim functions as a **premise** for a substantive conclusion in the same rhetorical unit—that is, when removing the process claim would break the argument for the substantive position. Otherwise, if the process evaluation is the terminal move of its own rhetorical unit (even when substantive arguments appear elsewhere in the turn), code as Process Evaluation.

*Operational test. If the process claim were removed, would the substantive argument still stand on its own grounds?* If yes, the process evaluation is terminal (Process Evaluation). If no—the substantive argument reads as a non-sequitur without the process claim—the process claim is a premise (Substantive Deployment).

*Positive markers (process claim as premise → Substantive Deployment).*

- Explicit inferential connectives linking process to substance: *entonces, por eso, así que, dado que, una vez que, ya que*, or analogous constructions. Example: *"ya oímos el criterio de la población, entonces sería desconocer todo ese amplio proceso de consulta, hay que llevarlo a votación."*
- The substantive conclusion grammatically depends on the process claim: removing the process clause produces a non-sequitur.
- The speaker uses the process claim to foreclose an alternative: *desconocer ese amplio proceso, sería contradictorio, no podemos ignorar*.

*Negative markers (process evaluation is terminal → Process Evaluation).*

- Process praise or critique appears in a discrete rhetorical unit (a gratitude move, an acknowledgment, a reflective aside, a closing affirmation) that is not grammatically linked to the substantive argument.
- The substantive argument is independently sustained—it has its own premises and would survive removal of the process evaluation.
- Process claims function as framing, mood-setting, or reflective celebration rather than as load-bearing warrant.

*Caveat on connectives.* The inferential connectives listed above are strong indicators but not sufficient conditions. A connective may appear without doing premise-work (e.g., *"por eso"* used as a discourse marker rather than as logical inference), and premise relations can hold without explicit connectives. Use the list as calibration, not as a pattern-match. The operational test—whether removing the process claim breaks the substantive argument—always takes precedence over the presence or absence of specific words.

*Worked examples (calibration only).* The operational test above, not the surface features of these examples, is the controlling rule.

*Terminal (Process Evaluation):* A speaker celebrates the consulta popular as transparent, courageous, and democratically valuable across several sentences of reflective gratitude, and elsewhere in the same turn advances a deferral argument grounded in independent reasons (the need to update the family code, the sensitivity of the topic). The praise is not syntactically linked to the deferral argument, and the deferral argument would stand without it. → Process Evaluation Positive.

*Premise (Substantive Deployment):* A speaker argues: *"Ya la población votó, ya oímos el criterio de la población. Entonces sería como desconocer todo ese amplio proceso de consulta. Hay que llevarlo a votación."* The process claim (*"ya oímos el criterio"*) functions as the premise for the substantive conclusion (*"hay que llevarlo a votación"*), linked by *"entonces."* Removing the process claim breaks the argument. → Substantive Deployment Positive.

*Process-about-process (Process Evaluation):* A speaker distinguishes supporting *consulta popular* from opposing *referéndum popular*, arguing that the two mechanisms play different institutional roles. The claim is about which consultation mechanism is appropriate—process evaluation about process—not a premise for a substantive constitutional provision. → Process Evaluation Mixed.

### A.2.6 Decision Procedure for Coding

For each speaker turn, follow this procedure:

**Step 1:** Read the entire speaker turn. Identify any references to formal consultation mechanisms (see §A.2.3).

**Step 2:** If no formal consultation references are present, output: `no\_invocations`.

**Step 3:** If references are present, segment into discrete invocations following the rules in §A.2.4.

**Step 4:** For each invocation, determine the rhetorical function by asking:

(a) Is the speaker evaluating the consultation process itself? → Process Evaluation.

(b) Is the speaker using consultation content to support, oppose, warrant, or elevate an issue (including procedural decisions like deferral)? → Substantive Deployment.

(c) Is there no discernible rhetorical function beyond factual reporting? → Descriptive Reference.

**Step 5:** Apply priority rules if multiple categories seem applicable.

**Step 6:** Assign valence where applicable (Process Evaluation, Substantive Deployment). Use the confidence field to indicate classifier certainty about the coding decision.

### A.2.7 Quick Reference: Categories and Valences

| Category | Core Question | Valence Options |
|---|---|---|
| Descriptive Reference | Is this pure factual reporting with no rhetorical function? | N/A |
| Process Evaluation | Is the speaker assessing the quality/legitimacy of consultation itself? | Positive / Negative / Mixed |
| Substantive Deployment | Is the speaker using consultation content to support, oppose, or elevate an issue? | Positive / Negative / Mixed |

### A.2.8 Guidance on Difficult Cases

**Blended references.** Speakers may blend formal consultation references with broader mandate claims (e.g., citing an IPN while also invoking "el estallido"). Code only the formal consultation reference. The broader framing is context but not a separate codable unit.

**Instrumental process evaluation.** A speaker who says "You rejected the most-signed IPN, which proves you don't respect the people" is using process evaluation instrumentally to oppose a substantive position. Code as Substantive Deployment (Negative), not Process Evaluation. A speaker who says "The consultation process was flawed because indigenous communities in remote areas couldn't participate" without connecting this to a substantive position is making Process Evaluation (Negative) the terminal move.

**Ambiguous mechanisms.** If a speaker references "lo que nos dijo la gente" or "las propuestas ciudadanas" without specifying a formal mechanism, examine context. If the reference is clearly tied to official consultation (e.g., within a discussion of IPN results), code it. If genuinely ambiguous between formal consultation and general mandate, do not code.

**Multiple mechanisms in one passage.** When a speaker cites multiple consultation mechanisms in a continuous passage serving the same rhetorical function (e.g., "We received 119 audiencias, 27 indigenous initiatives, and 4 IPNs"), code as one invocation with mechanisms listed together. The segmentation rule "one invocation per distinct referent" applies when mechanisms serve distinct rhetorical functions or are separated by intervening material. If the same passage deploys one mechanism positively and another negatively, code as separate invocations.

**Reported invocations.** Code only direct invocations by the speaker. If Speaker A reports that Speaker B invoked consultation (e.g., "My colleague claimed the IPN supports her position"), this is not a codable invocation by Speaker A. Speaker A is describing another's representative claim, not making one. However, if Speaker A then evaluates or deploys that reported content (e.g., "My colleague claimed the IPN supports her position, but she misrepresents what citizens actually proposed"), code Speaker A's evaluation or deployment.

### A.2.9 Lexical Cues for Identification

This section provides lexical cues to assist in identifying candidate passages for coding. The terms below are strong indicators of formal consultation references, but they do not constitute an exhaustive list. Flag speaker turns containing these terms AND any other references to formal consultation mechanisms as defined in §A.2.3, even if phrased differently.

Use this list as calibration, not constraint. If a passage clearly references a formal consultation mechanism using language not on this list (e.g., "lo que recogimos en terreno" for semanas territoriales), it should still be flagged and coded. Conversely, the presence of a term from this list does not guarantee a codable invocation—the reference must meet the definitional criteria in §A.2.3.

**Chile—n-gram search terms.** *Iniciativas Populares de Norma (IPNs):* iniciativa popular, iniciativas populares, iniciativa ciudadana, IPN.

*Named initiatives and sponsoring organizations (non-exhaustive).* Presence of a name from the list below does not itself establish a codable invocation. The reference must meet §A.2.3 criteria—the initiative must be invoked as formal consultation input, not mentioned incidentally (e.g., a speaker naming their own sponsoring organization, a historical aside, or a passing reference to the initiative's existence without deploying its content or evaluating the process that produced it):

> Con mi plata NO; Será Ley; Banco Central Autónomo; Derecho a la Vida; Siempre por la Vida; Acción Educar; #NOSONMUEBLES; Vegetarianos Hoy; Compromiso Minero; NO+AFP; Nacionalización del cobre; Soy del campo; Nos Importan; Pobladoras y Pobladores; Multigremial Nacional de Emprendedores; Sociedad Civil por la Acción Climática; REDOFEM; Sistema único de salud; Consorcio de Universidades del Estado; #iniciativapopularfeminista; SacaLaVoz; Sujetos no objetos; Compromiso con la niñez; Vive Chile Rural; SOFOFA; Emprende Libre; TEAmoMas; zonas de sacrificio; no más TAG; Colectivo Nacional por la Discapacidad; Chile Mejor sin TLC; Revolución Ciclista; D.U.R.A.S.; presos políticos de la rebelión; Consentimiento médico informado; Estado laico no laicista; CUT; central unitaria de trabajadores; CTC; Renacionalización del cobre; CONFUSAM; fin del sistema de SENAME; Objeción de conciencia; Es mi plata y punto; G9; ABOFEM; muerte digna; ANAMURI; Derecho a la alimentación; FLADEM; FENPRUSS; Vertebral; CIMUNIDIS; ISP; 15.000 corazones por la tierra; derechos de la naturaleza; Derecho a ser club y ser hincha; Derecho a los cuidados; pueblo tribal afrodescendiente; Cannabis a la Constitución; Educación Libre y Diversa; Primero las victimas; Por el agua los derechos de la naturaleza y los glaciares.

*Audiencias públicas:* audiencia pública, audiencias públicas, la audiencia, una audiencia.

*Consulta indígena:* consulta indígena, consultas indígenas, iniciativas constituyentes indígenas, iniciativas indígenas.

*Cabildos and territorial mechanisms:* cabildo, cabildos autoconvocados, encuentros autoconvocados, semana territorial, semanas territoriales, jornada nacional, jornadas nacionales, encuentros territoriales.

*Participation infrastructure:* participación ciudadana, participación popular, Secretaría de Participación, Secretaría de Participación Popular, Secretaría de Participación y Consulta Indígena.

**Cuba—n-gram search terms.** *Note:* In the Cuban context, the consulta popular is the dominant frame for the entire constitutional debate. Terms that would be generic in Chile (e.g., "opiniones," "criterios," "planteamientos") frequently refer specifically to consultation results when used by speakers discussing the constitutional process. Code these as consultation references when context indicates they refer to input gathered through the formal consulta popular.

*Consulta popular (direct references):* consulta popular, proceso de consulta, la consulta, consulta del pueblo, discusión popular.

*Consultation results (terms referring to input received):* opiniones, criterios, planteamientos, propuestas, propuestas de la población, propuestas ciudadanas, propuestas populares, criterios de la población, planteamientos de la población, puntos de vista.

*Quantitative references.* Aggregate statistics citing numbers of opinions, percentages, or meetings (e.g., "51.414 opiniones," "el 6,56% del total," "40.420 reuniones"). These numerical references to consultation data are strong indicators.

*Process and legitimacy language:* ejercicio democrático, sabiduría popular, sentir del pueblo, voluntad del pueblo (when linked to the consultation process).

**Context-dependent terms.** The following terms require contextual judgment. In Chile, they typically refer to delegate proposals or procedural matters and should not be used as primary search terms. In Cuba, they frequently refer to consultation results and should be coded when context indicates reference to the consulta popular:

propuestas, opiniones, criterios, planteamientos, reuniones, aportes, modificaciones.

Contextual indicators that these terms refer to formal consultation: mention of statistics or percentages, explicit linkage to "la consulta" or "el pueblo," discussion of what the commission received or processed, framing as input that was accepted, rejected, or considered.

*Threshold caveat (Cuba).* The lower threshold for context-dependent terms in Cuba does not

override §A.2.8. Absent at least one of the contextual indicators above, do not code. Presence of the term alone is insufficient—a speaker saying "en mi opinión" or "tengo el criterio de que" without contextual linkage to the consulta popular is not a codable invocation.

### A.2.10 Codebook Version History

The codebook went through 10 numbered revisions during development. The full changelog is released with the replication package (§A.9). The three revisions cross-referenced elsewhere in this appendix are summarized briefly below.

**Version 2.5.** Removed theoretical framing from the codebook overview. The framing risked priming the classifier toward reading neutral references as rhetorically active, compromising the textual diagnosability of Descriptive Reference as the residual category.

**Version 2.6.** Tightened the Substantive Deployment valence anchor so that valence tracks the speaker's orientation toward the provision being argued, not toward the cited input itself. Added a "When to use Mixed" rule and tightened confidence calibration.

**Version 2.7.** Required `text_excerpt` to be a verbatim substring of the speaker turn; enforced via the pipeline's `unlocatable_excerpts` log column. A complementary sentence-boundary normalizer extends each excerpt to its full enclosing sentence(s) at the CSV-flattening stage.

## A.3 Prompts and output schema

The system prompt is constructed from a fixed template, the country-specific guidance for the corpus being processed, and the full codebook (§A.2). The codebook block is sent under `cache\_control:\{"type":"ephemeral"\}` so it is paid once per cache window and reused across all turns in a country. Each speaker turn arrives as a user message containing `session\_id`, `row\_id`, country, and the verbatim text.

### A.3.1 Canonical system prompt

The canonical system prompt template is reproduced verbatim below. `\{country\}`, `\{country\_specific\_guidance\}`, and `\{codebook\_text\}` are substituted at runtime.

You are a research assistant coding constitutional assembly transcripts for references to formal public consultation mechanisms. Your task is to identify and classify how constitutional drafters invoke public input in plenary debates.

COUNTRY CONTEXT: {country}

{country_specific_guidance}

CODEBOOK:

{codebook_text}

CRITICAL INSTRUCTIONS:

- Follow the decision procedure in Section 6 exactly.
- If there are no references to formal consultation mechanisms, output exactly: `\{"has\_invocations":false,"invocations":[]\}`
- Output ONLY a valid JSON object with two fields: `has\_invocations` (boolean) and `invocations` (array of invocation objects per the Section 7 schema).
- Each invocation object must contain: `invocation\_id`, `mechanism`, `text\_excerpt`, `rhetorical\_function`, `valence`, `confidence`, `rationale`.
- Do not include `session\_id`, `row\_id`, speaker name, or metadata — the script adds those.
- Do not include any text before or after the JSON object.
- Do not wrap the JSON in markdown code fences or any other formatting.
- Err on the side of `no\_invocations` for genuinely ambiguous cases (Section 9.3).
- Before emitting multiple invocations within a single speaker turn, apply the §4.2.2 operational test: could the passage between the two consultation references be removed without changing what the speaker is arguing about? If yes, code as one invocation, not two. Default to fewer invocations when uncertain.
- Before coding a process-evaluation passage as Substantive Deployment under Rule 1, apply the §5.2.1 premise test: if the process claim were removed, would the substantive argument still stand on its own grounds? If yes, code as Process Evaluation (the process claim is terminal). If no, code as Substantive Deployment (the process claim is a premise). Default to Process Evaluation when uncertain.
- `text\_excerpt` must be a VERBATIM substring of the speaker's turn. Copy the passage exactly — do not paraphrase, do not rewrite, do not normalize disfluencies, and do not silently stitch together

non-contiguous portions of the turn as if they were continuous. If you need to elide material in the middle of a contiguous passage, mark the elision explicitly with "…" within the single excerpt. If the invocation spans non-contiguous portions of the turn, code them as separate invocations per §4.1, not as one stitched excerpt. Preserve the original punctuation, spacing, and wording exactly as it appears in the turn.

### A.3.2 Country-specific guidance

**Cuba.** You are coding a CUBAN constitutional assembly transcript (2018-2019 National Assembly debates on the 2019 Constitution). The dominant formal consultation mechanism is the CONSULTA POPULAR — a mass nationwide consultation process. Terms such as 'opiniones', 'criterios', 'planteamientos', 'propuestas', and 'reuniones' frequently refer to results of the consulta popular when used in the context of discussing the constitutional draft. Aggregate statistics (e.g., '51,414 opiniones', '40,420 reuniones') are strong indicators of a formal consultation reference. Apply the Cuba-specific guidance in Appendix A.2 and the context-dependent terms guidance in Appendix A.3.

**Chile.** You are coding a CHILEAN constitutional convention transcript (2021-2022 Convención Constitucional). Formal mechanisms include: Iniciativas Populares de Norma (IPNs), audiencias públicas, consulta indígena, cabildos autoconvocados, and semanas territoriales. Context-dependent terms like 'propuestas', 'opiniones', and 'criterios' typically refer to delegate proposals or procedural matters, NOT formal consultation references, unless explicitly linked to a named formal mechanism. Apply the Chile-specific guidance in Appendix A.1 and the context-dependent terms guidance in Appendix A.3.

### A.3.3 Paraphrased system prompt

The paraphrased prompt restates the same task with different surface wording while preserving all definitional content. It is used in Run 4 (Table 1) to assess inter-prompt reliability ($\alpha_{\text{inter-prompt}}$ in §A.4.4). The country-specific guidance and codebook block (the latter sent as cached prefix) are identical to the canonical run. Only the task framing and instruction list change:

You are an expert research coder analyzing constitutional assembly transcripts. Your task is to identify

passages where drafters reference formal public consultation mechanisms and classify each reference according to the codebook provided below.

COUNTRY CONTEXT: {country}

{country_specific_guidance}

CODEBOOK:

{codebook_text}

RULES:

- Apply the Section 6 decision procedure precisely as written.
- When a speaker turn contains no reference to any formal consultation mechanism, return exactly: `\{"has\_invocations":false,"invocations":[]\}`
- Your response must be a single valid JSON object containing `has\_invocations` (boolean) and `invocations` (an array of invocation objects matching the Section 7 schema).
- Every invocation object must include these fields: `invocation\_id`, `mechanism`, `text\_excerpt`, `rhetorical\_function`, `valence`, `confidence`, `rationale`.
- Do not add `session\_id`, `row\_id`, speaker name, or any other metadata — those are appended by the pipeline.
- Return only the JSON object. No surrounding text, no markdown fences, no additional formatting.
- For genuinely ambiguous passages, prefer `no\_invocations` over forced classification (Section 9.3).
- When considering whether to split a single speaker turn into multiple invocations, use the §4.2.2 operational test: ask whether the text between two candidate consultation references could be deleted without altering the speaker's argument. If so, treat them as a single invocation. When in doubt, code fewer invocations rather than more.
- When a passage contains both process commentary and substantive content, apply the §5.2.1 premise test: ask whether the substantive argument would hold even if the process claim were deleted. If the argument stands alone, the process claim is terminal — code as Process Evaluation. If the argument depends on the process claim, code as Substantive Deployment. When in doubt, prefer Process Evaluation.
- The `text\_excerpt` field must contain a VERBATIM copy of the relevant passage from the speaker's turn. Do not paraphrase, rephrase, clean up disfluencies, or concatenate non-adjacent text

segments into a single excerpt. If you need to omit material within a contiguous passage, indicate the gap with "…" inside that excerpt. If the reference spans non-contiguous parts of the turn, code each part as a separate invocation per §4.1. Preserve original punctuation, spacing, and wording exactly.

### A.3.4 Output schema

For each speaker turn, the classifier outputs a JSON object with the following structure:

```
{
  "has_invocations": true/false,
  "invocations": [
    {
      "invocation_id": 1,
      "mechanism": "[specific mechanism referenced]",
      "text_excerpt": "[relevant portion of speaker text]",
      "rhetorical_function": "[category name]",
      "valence": "[valence or null if not applicable]",
      "confidence": "high/medium/low",
      "rationale": "[brief explanation of coding decision]"
    }
  ]
}
```

If no invocations are present, the output is:

```
{ "has_invocations": false, "invocations": [] }
```

**Field constraints (enforced post-hoc).**

- `rhetorical\_function` ∈ {Descriptive Reference, Process Evaluation, Substantive Deployment}.
- `valence` ∈ {Positive, Negative, Mixed, null}; null only when `rhetorical\_function` is Descriptive Reference.

- `confidence` $\in$ {high, medium, low}.
- `text\_excerpt` must be a verbatim substring of the speaker turn (whitespace- and punctuation-normalized). Each turn's count of unlocatable excerpts is logged to `processing\_log.csv` via the `unlocatable\_excerpts` field; these are flagged for review and excluded from the analytic CSV if they cannot be reconciled.

**Text excerpt convention.** The excerpt must include enough context to make the rhetorical function legible without reference to the full speaker turn. `text\_excerpt` must be a VERBATIM substring of the speaker's turn: copy the passage exactly, preserving original punctuation, spacing, and wording. Do not paraphrase, rewrite, normalize disfluencies, or silently stitch together non-contiguous portions of the turn. If an elision within a single contiguous passage is necessary, mark it explicitly with "…" inside the excerpt. Invocations that span non-contiguous portions of the turn must be coded as separate invocations per §A.2.4, not as one stitched excerpt.

**Confidence levels.** High = the §A.2.3 mechanism is clearly referenced AND the decision procedure yields one rhetorical function (and valence) unambiguously. Medium = one of those two conditions is uncertain but not both (e.g., mechanism is clear but function is contestable, or function is clear but the reference is borderline formal-vs-general). Low = both conditions are uncertain, or the coder would find it hard to defend the assigned category against a reasonable alternative. Medium is not a safe default—if the coding decision is straightforwardly supported by the rules, assign High.

## A.4 Validation

The validation protocol has four layers, each targeting a distinct failure mode: schema validity (§A.4.1), classification precision against human review on a stratified sample (§A.4.2), recall against an independent lexical floor (§A.4.3), and reliability across repeated and paraphrased runs (§A.4.4). Layers 2 and 3 surfaced specific systematic miscoding patterns, all corrected by hand-coded overrides folded into the analytic dataset; the audit notes for those investigations are reproduced in §A.4.5.

### A.4.1 Schema validity and run integrity

Across all v2.7 runs (canonical Run 1 on the full corpus, canonical Runs 2–3 and paraphrased Run 4 on the validation subsample), the pipeline processed roughly 56,700 speaker turns. Every output is checked against the JSON schema in §A.3.4 (field presence, allowed values, valence/function compatibility) and against the verbatim-substring requirement on `text\_excerpt`. Schema-invalid responses are excluded from the analytic CSV and recorded with status `error` in `processing\_log.csv`; only one such row appears across all v2.7 logs. The `unlocatable\_excerpts` field added in v2.7 (see §A.2.10) flags excerpts the verbatim-substring check cannot reconcile against the source turn after whitespace and punctuation normalization. Across the full Cuba and Chile v2.7 runs, two such cases were flagged (both in long Acosta turns) and were resolved by re-extracting from the source.

### A.4.2 Stratified human audit

A stratified blind audit hand-coded roughly 25 invocations from each populated country $\times$ function-valence cell: 150 from Chile across 6 cells and 93 from Cuba across 4 cells (Cuba's grid populates only 4 of 7 cells; see main text §5.2). All audit subjects were drawn from canonical Run 1. Headline metrics:

| **Metric** | **Chile (n=150)** | **Cuba (n=93)** |
|---|---|---|
| Detection precision (true positives / sample) | 96.7% (145/150) | 98.9% (92/93) |
| Full classification agreement (function + valence) | 92.0% (138/150) | 94.6% (88/93) |
| Full classification agreement \| true positive | 95.2% (138/145) | 95.7% (88/92) |
| Function agreement \| true positive | 96.6% (140/145) | 95.7% (88/92) |
| Valence agreement \| true positive (valenced strata) | 96.7% (n=121) | 100.0% (n=68) |

Per-stratum accuracy, false-positive counts, and disagreement decompositions are in Tables 2 and 3.

**Sample-level disagreement patterns.** Chile's 12 disagreements break down as 5 false positives, 4 function-only errors (with valence preserved), 2 valence-only errors, and 1 case where both fields

Table 2: Chile stratified audit: per-stratum accuracy (n=150).

| Stratum | n | Agree | FP | Function wrong | Valence-only wrong | Fn + val wrong | Acc. |
|---|---|---|---|---|---|---|---|
| Descriptive Reference | 25 | 23 | 1 | 1 | 0 | 0 | 92.0% |
| Process Evaluation / Mixed | 25 | 22 | 1 | 0 | 1 | 1 | 88.0% |
| Process Evaluation / Negative | 25 | 22 | 0 | 2 | 1 | 0 | 88.0% |
| Process Evaluation / Positive | 25 | 24 | 0 | 1 | 0 | 0 | 96.0% |
| Substantive Deployment / Negative | 25 | 24 | 1 | 0 | 0 | 0 | 96.0% |
| Substantive Deployment / Positive | 25 | 23 | 2 | 0 | 0 | 0 | 92.0% |

Table 3: Cuba stratified audit: per-stratum accuracy (n=93). Process Evaluation / Negative, Process Evaluation / Mixed, and Substantive Deployment / Mixed cells were unpopulated in Run 1 and not sampled.

| Stratum | n | Agree | FP | Function wrong | Valence-only wrong | Fn + val wrong | Acc. |
|---|---|---|---|---|---|---|---|
| Descriptive Reference | 25 | 22 | 1 | 2 | 0 | 0 | 88.0% |
| Process Evaluation / Positive | 25 | 24 | 0 | 1 | 0 | 0 | 96.0% |
| Substantive Deployment / Negative | 18 | 18 | 0 | 0 | 0 | 0 | 100.0% |
| Substantive Deployment / Positive | 25 | 24 | 0 | 1 | 0 | 0 | 96.0% |

differed. All 5 false positives are scope-diffusion errors on *consulta indígena*—the model coded references to indigenous consultation as an external legal institution, a future norm-design standard, or a historical regulatory regime as if they were references to the Convention's own CI process. This pattern motivated the systematic CI-scope audit summarized in Table 6, which flagged 5 additional false positives across the 71 unreviewed CI invocations. Cuba's 5 disagreements are 1 false positive, 4 function-only errors, and 0 valence errors; valence agreement on the valenced strata was 100%.

### A.4.3 Lexical recall floor

To bound false negatives, the lexical recall floor (`scripts/validation/lexical\_recall\_floor.py`) scans every predicted-negative turn for the lexical cues in §A.2.9. Any hit is a candidate the classifier may have missed. Because semantic invocations without any of these lexical cues cannot be detected by this method, the audit bounds recall error from below; it does not estimate it directly.

| | Chile | Cuba |
|---|---|---|
| Predicted-negative turns scanned | 22,501 | 1,092 |
| Turns flagged by at least one cue | 1,233 (5.5%) | 7 |
| of which high-precision-cue hits reviewed | 288 | 7 |
| Confirmed false negatives | 4 | 1 |
| Implied keyword-bounded miss rate | 0.018% | 0.09% |

The Chile audit prioritized turns hitting high-precision mechanism terms (iniciativa popular / IPN, audiencias públicas, consulta indígena, cabildos autoconvocados, semanas territoriales, encuentros / jornadas territoriales, Secretaría de Participación). Generic context-dependent terms (*propuestas, criterios, opiniones, modificaciones*) were not individually reviewed given their known high noise rate in Chile (codebook §A.2.9). The 4 Chile false negatives span three different mechanisms (semanas territoriales / cabildos, audiencias públicas, consulta indígena), so there is no evidence of a mechanism-specific recall blind spot. The 1 Cuba false negative was a Descriptive Reference mention of opinions generated during the consulta popular. All 5 cases were added as upgrades to the human-overrides files and folded into the analytic dataset; full case notes appear in §A.4.5.

### A.4.4 Inter-run and inter-prompt reliability

Reliability was assessed at three levels using three canonical-prompt runs (`run\_1`, `run\_2`, `run\_3`; temperature 0) for self-consistency and one paraphrased-prompt run (`run\_4`) for prompt sensitivity. Inter-run agreement was computed on Cuba's full corpus (265 turns) and on a 13-session, 2,944-turn Chile subsample that oversampled invocation-rich sessions (41% of Chile's positive turns) so classification-level $\alpha$ has adequate denominators. Tables 4 and 5 report Krippendorff's $\alpha$ at each level for each corpus.

All four runs produce $\alpha$ above the 0.91 threshold reported in the main text, comfortably exceeding the 0.80 threshold conventionally used for reliable agreement (Krippendorff 2019). The level-1 detection $\alpha$ in Chile is depressed slightly by the prevalence effect: the positive-turn base

Table 4: Chile inter-run agreement (13-session subsample, 2,944 turns).

| | **Self-consistency (runs 1–3)** | **All runs (1–4, incl. paraphrased)** |
|---|---|---|
| *Level 1: Binary detection* | | |
| Percent agreement | 99.8% | 99.3% |
| Krippendorff's $\alpha$ | 0.990 | 0.970 |
| *Level 2: Invocation count \| detection* | | |
| Percent agreement | 89.7% (n=174) | 81.5% (n=173) |
| Krippendorff's $\alpha$ (interval) | 0.921 | 0.873 |
| *Level 3: Classification \| count match* | | |
| Function: % agreement / $\alpha$ | 95.6% / 0.948 | 90.8% / 0.911 |
| Valence: % agreement / $\alpha$ | 95.6% / 0.955 | 93.4% / 0.942 |

Table 5: Cuba inter-run agreement (full corpus, 265 turns).

| | **Self-consistency (runs 1–3)** | **All runs (1–4, incl. paraphrased)** |
|---|---|---|
| *Level 1: Binary detection* | | |
| Percent agreement | 100.0% | 99.6% |
| Krippendorff's $\alpha$ | 1.000 | 0.994 |
| *Level 2: Invocation count \| detection* | | |
| Percent agreement | 87.5% (n=56) | 76.8% (n=56) |
| Krippendorff's $\alpha$ (interval) | 0.984 | 0.977 |
| *Level 3: Classification \| count match* | | |
| Function: % agreement / $\alpha$ | 99.1% / 0.988 | 95.6% / 0.956 |
| Valence: % agreement / $\alpha$ | 99.1% / 0.988 | 96.7% / 0.966 |

rate is 6.0%, and $\alpha$ is sensitive to base-rate imbalance even at near-perfect agreement (the disagreement count is just 20 turns out of 2,944). The principal source of cross-run drift is at level 2 (invocation count), where the paraphrased prompt occasionally splits or merges where the canonical prompt does not; level 3 classification, conditional on count match, holds essentially constant across canonical and paraphrased runs.

### A.4.5 Targeted audits

Disagreement patterns surfaced by the stratified sample (§A.4.2) and the lexical recall floor (§A.4.3) motivated five targeted audits, summarized in Table 6. All corrections were appended to country-specific human-overrides files and folded into the analytic dataset before statistical analysis. Validation metrics in §A.4.2 are sample-only and do not include these out-of-sample corrections.

Table 6: Targeted audits and corrections folded into the analytic dataset.

| Date | Audit | Scope | Corrections recorded |
|---|---|---|---|
| 2026-04-14 | Cuba lexical recall floor | 1,092 predicted-negative turns | 1 false negative (+1 Descriptive Reference) |
| 2026-04-17 | Cuba SD valence anchoring | 94 SD cases (all) | 4 valence flips (Pos→Neg); 1 exclusion (historical-process reference) |
| 2026-04-17 | Chile SD valence anchoring | 231 SD cases (all) | 2 valence flips (Pos→Neg) |
| 2026-04-21 | Chile consulta indígena scope | 99 CI invocations (all) | 5 false positives (scope diffusion: external CI / future norm / historical regulatory regime); re-coded `no_invocations` |
| 2026-04-21 | Chile lexical recall floor | 22,501 predicted-negative turns; 1,233 lexical hits; 288 high-precision-cue hits manually reviewed | 4 false negatives (+2 SD/Positive, +1 SD/Negative, +1 Process Evaluation/Positive) |

*Notes:* Cuba corrections appended to `data/validation/cuba/human_overrides_cuba.csv` (6 rows total). Chile corrections appended to `data/validation/chile/human_overrides_chile.csv` (11 rows total). The Chile SD valence audit identified three errors but recorded two; the third (Hoppe Espoz, Sesión 80) was not folded into the analytic dataset. Recall-floor miss rates implied by the lexical audits are reported in §A.4.3.

Per-case `uids`, source-text excerpts, and rationales are recorded in the `human_notes` column of the overrides files.

**Chile valence-anchoring audit (2026-04-17).** *Context.* Following the Cuba valence anchoring audit (same date), the same error pattern was checked across all Chile SD cases. The error: the model anchors SD valence to the speaker's orientation toward the consultation input rather than toward the provision under debate. Chile was expected to have a higher error rate due to its more adversarial deliberative context, where delegates more frequently deploy consultation mechanisms to argue against provisions.

*Scope of audit.* All 231 Substantive Deployment cases in Chile run_1 were reviewed: 176 SD/Positive and 55 SD/Negative. Keyword-based screening flagged 43 SD/Positive cases for manual review; all 55 SD/Negative cases were reviewed in full. Source transcripts were read

for ambiguous cases.

*Finding: same input-anchoring pattern, lower prevalence.* Three valence errors were identified, all SD/Positive that should be SD/Negative. In each case the speaker embraces consultation input but deploys it to argue against the provision as drafted. The model anchored to the speaker's orientation toward the input (Positive) rather than toward the provision (Negative). Contrary to the initial hypothesis, Chile's error rate (1.3%) is lower than Cuba's (4.3%), possibly because the adversarial context makes the model more likely to code opposition as Negative in the first place.

*Errors identified.* Three valence errors (Positive → Negative):

- *Sesión_80_Plenario_r500_i1* (Vanessa Hoppe Espoz): deploys consulta indígena as warrant for rejecting articles 24 bis, indicación 40, and artículo 26, arguing indigenous justice regulation should be deferred to law rather than limited in the constitutional text.
- *Sesión_91_Plenario_r98_i1* (Mauricio Daza Carrasco): deploys audiencias públicas held by Commission 6 to argue that the Plenary should reject articles 11 bis and 12 from the current report in favor of the Commission 6 proposals.
- *Sesión_106_Plenario_r91_i1* (Natalia Henríquez Carreño): deploys semanas territoriales input to argue against Article 2 of the transitional provisions.

*Indicaciones convention.* Chile's Convention uses indicaciones (formal amendments) as a standard legislative tool. When a speaker deploys consultation input to support an indicación, the indicación itself is treated as the provision under consideration for valence coding purposes. This convention avoids systematic miscoding of amendment-supporting speakers as Negative.

*Combined error rate.* Across both countries: 7 valence errors out of 325 total SD cases (2.2%). All errors are the same type: SD/Positive coded when the speaker deploys consultation input to argue against the provision as drafted (should be SD/Negative).

**Chile consulta-indígena scope audit (2026-04-21).** *Context.* The stratified validation sample (`sample\_chile.csv`) flagged 5 false positives among 28 consulta indígena cases (17.9%). All five traced to scope diffusion—speakers using "consulta indígena" to refer to something other

than the Chilean Constitutional Convention's own CI process. Three sub-patterns: (a) CI as an external legal institution or ILO 169 right, (b) CI pertaining to a specific provision's design or future application rather than the Convention's completed CI, and (c) prospective CI embedded in a norm being debated.

*Scope of audit.* All 99 Chile run_1 invocations tagged with `mechanism='Consultaindgena'` were reviewed (28 already in the sample; the remaining 71 examined systematically in batches).

*Finding: five additional false positives.* The audit identified 5 FPs in the 71 unreviewed rows (7.0%). All 5 fit one of the three sub-patterns. Combined with the 5 sample FPs, run_1 CI precision is 89.9% (89/99). The 5 audit FPs (re-coded `no\_invocations`):

- *Sesión_55_Plenario_r136_i1* (John Smok): mechanism misclassification—speaker cites 'iniciativa constituyente indígena 257', a distinct participation mechanism from the Convention's CI.
- *Sesión_66_Plenario_r157_i2* (Ericka Portilla Barrios): scope diffusion (ILO 169 right)—CI invoked re: lithium extraction at Salar de Maricunga as an external administrative consultation right.
- *Sesión_91_Plenario_r136_i1* (Fuad Chahín Valenzuela): scope diffusion (constitutional design standard)—CI treated as the ILO 169 standard to be enshrined in draft article 16.
- *Sesión_92_Plenario_r175_i2* (Wilfredo Bacián Delgado): scope diffusion (historical legal framework)—speaker critiques the pre-Convention regulatory regime governing administrative CIs (decretos 66/40). Excluded per §A.2.3.
- *Sesión_107_Plenario_r124_i2* (Margarita Vargas López): prospective reference—the norm under debate embeds a future CI for Autonomías Territoriales Indígenas; the CI is a downstream requirement of the provision, not a completed consultation being cited.

*Relationship to validation sample.* The sample FP rate (5/28 = 17.9%) over-represented the true audit rate (10/99 = 10.1%) because stratified sampling drew from CI-keyword-matched rows

and concentrated borderline scope cases.

**Chile lexical-recall audit (2026-04-21).** *Context.* Two prior Chile audits (SD valence anchoring, CI scope) measured precision on predicted-positive turns, leaving recall unmeasured. The lexical recall floor scans every predicted-negative turn for the codebook's lexical cues; any hit is a potential false negative.

*Scope of audit.* All 22,501 Chile run_1 predicted-negative turns were scanned. 1,233 turns contained at least one cue (5.5% raw hit rate). Manual review focused on the 288 turns hitting high-precision mechanism terms (iniciativa popular / IPN, audiencias públicas, consulta indígena, cabildos autoconvocados, semanas territoriales, encuentros / jornadas territoriales, Secretaría de Participación). Generic context-dependent terms (*propuestas, criterios, opiniones, modificaciones*) were not individually reviewed given their known high noise rate in Chile (codebook §A.2.9).

*Finding: four false negatives identified.* Four recurring sources of lexical-floor false positives dominate: (a) 'Cabildo' as a Valparaíso Region town name; (b) prospective constitutional design (draft articles enshrining future IPNs, CI, referenda, or audiencias as mechanisms), which is out of scope per §A.2.3; (c) procedural / infrastructure discussions; (d) rhetorical or administrative mentions unrelated to consultation-content deployment. Four genuine missed invocations were confirmed:

- *Sesión_58_Plenario_r196_i1* (Yarela Gómez Sánchez): cites the more than 25 cabildos she organized across 4 semanas territoriales in 22 Aysén localities. Substantive Deployment / Positive; mechanism Semanas territoriales / Cabildos autoconvocados.

- *Sesión_61_Plenario_r182_i1* (Javier Fuchslocher Baeza): evaluates the Biobío audiencias públicas experience positively and uses that evaluation to argue in favor of the Mesa proposal to plan a Pleno outreach trip to Antofagasta. Process Evaluation / Positive; mechanism Audiencias públicas.

- *Sesión_93_Plenario_r234_i1* (Benito Baranda Ferrán): cites expert testimony from audiencias públicas in the Comisión de Derechos Fundamentales as auxiliary warrant for rights-specification.

Substantive Deployment / Positive; mechanism Audiencias públicas.

- *Sesión_94_Plenario_r46_i1* (Carol Bown Sepúlveda): briefly evaluates the Convention's participation processes as having had *'escasa, si no nula participación indígena'* and deploys that evaluation as a premise for opposing article 81 of the Sistema de Justicia report. Substantive Deployment / Negative; mechanism Consulta indígena.

*Implication for recall.* Four confirmed misses across 22,501 predicted negatives yields a lexical-floor miss rate of approximately 0.018%. True recall error is certainly higher since semantic invocations without lexical cues cannot be caught by this method. The audit rules out a catastrophic recall failure on turns the classifier did see lexical triggers in. The miss pattern spans three different mechanisms (semanas territoriales / cabildos, audiencias públicas, consulta indígena), so there is no evidence of a mechanism-specific recall blind spot.

## A.5 Corpora and speaker metadata

### A.5.1 Chile corpus

The Chile corpus comprises 80 plenary session transcripts (Sesión 31–110), archived by the Library of the National Congress of Chile and obtained as Word documents. Sesiones 1–30 were excluded because the Convention's General Regulations were not approved until 13 October 2021 (Sesión 30); substantive constitutional debates begin with Sesión 31. Across the 80 sessions there are 22,557 speaker turns and approximately 2.36 million words. Each session file was parsed into one row per speaker turn, with non-speech text (vote recordings, procedural roll calls, time stamps) removed during pre-processing. Each row carries the row-level text plus speaker-level metadata merged from external sources.

### A.5.2 Cuba corpus

The Cuba corpus consists of three sessions of the National Assembly held December 20–22, 2018—the post-consultation debate of the 2019 constitutional draft. No official transcripts were ever published. The sessions were broadcast as Facebook livestreams by *Cubadebate*, a state-run media platform. The video was downloaded, audio extracted, and transcribed using `WhisperX`

(Bain et al. 2023), an automatic speech recognition system for long-form audio with built-in speaker diarization (Radford et al. 2023). Transcripts were checked against the audio for speaker identification and segmentation. The combined corpus contains 265 speaker turns and 73,276 words across the three sessions. The December 21 session features a roughly 30,000-character monologue by Homero Acosta Álvarez (Secretary of the Council of State) walking the Assembly through the draft article-by-article; this turn was split into multiple rows during pre-processing to accommodate Excel's per-cell character limit, but all sub-rows share session and speaker identifiers and were re-joined for any analysis at the speaker level.

### A.5.3 Speaker metadata

**Chile.** Speaker-level data are merged from two sources. Demographic and biographical fields (district, region, sex, date of birth, education level, occupation, marital status, electoral list, party, collective affiliation) come from (Rozas et al. 2023). Continuous ideology scores come from (Fábrega 2022), which estimates a one-dimensional left-right position from delegate roll-call votes during the constitutional process; the score is bounded on $[-1, +1]$, with negative values denoting left and positive values right. The score is published with an associated standard deviation; both fields appear in the analytic dataset (`ideology1`, `sd1`, with a secondary specification `ideology2`, `sd2`). Four binary commission-status indicators code each delegate's participation in the Convention's provisional, thematic, and miscellaneous (voluntary) commissions, and service on the executive committee (Mesa Directiva).

**Cuba.** Demographic and biographical fields (sex, age, occupation, education level, municipality, province, vote percentage, mass-organization affiliations) come from Proyecto Inventario, a research initiative cataloging the Cuban political elite. Three binary indicators code each deputy's membership in the PCC Central Committee (`central\_committee`), the Politburo (`politburo`), and the 33-member Drafting Commission (`draft\_comm`). Cuba lacks a comparable continuous ideology measure; the Cuban analyses in the main text use these elite-status indicators as the relevant institutional cleavage. Delegates whose names did not match the Proyecto Inventario

roster (3 of 70 floor speakers) are excluded from any analysis that joins on metadata.

### A.5.4 Bloc and Drafting-Commission coding

**Chile—Vamos por Chile.** The bloc indicator `is\_vpc` equals 1 for any delegate whose electoral list (`elec\_list`) is exactly *Vamos por Chile*, the right-wing list that won 37 of 155 seats in the May 2021 election. This is the sole formal political bloc with a coherent ideological position throughout the Convention's life; the remaining 118 seats fragment across Apruebo Dignidad (left), Lista del Apruebo (center-left), La Lista del Pueblo (left-populist independents), Independientes por la Nueva Constitución (centrists), Asamblea Popular Constituyente (left independents), the 17 reserved-seat indigenous delegates, and several smaller regional or movement-based lists. Because the analysis tests whether VpC's behavior diverges from the rest of the Convention, the comparison is binary (VpC vs. non-VpC) rather than across the full list partition. Robustness to alternative bloc encodings is discussed in §A.6.4.

**Cuba—Drafting Commission.** The Drafting Commission was a 33-member committee constituted by the National Assembly in June 2018 to prepare the constitutional draft. Membership was published in the official ANPP records and is cross-validated against Proyecto Inventario. The indicator `draft\_comm` equals 1 for the 33 commissioners. The two ancillary elite indicators (`central\_committee`, `politburo`) are coded from PCC organizational charts as of December 2018.

## A.6 Robustness checks

### A.6.1 Permutation and bootstrap tests

The headline statistical claims in the main text are reinforced by resampling. All resampling uses 10,000 draws; seeds are fixed (`20260421` for Chile, `20260420` for Cuba) and reproducible from `scripts/analysis/\{country\}\_statistics.py`.

**Chile—VpC concentration of Negatives.** The 460-invocation valenced pool (Positives + Negatives) contains 176 Negatives, of which VpC produces 122. Under the null of label exchangeability across speakers in the valenced pool, the expected VpC Negative count is $\approx 50.5$ (VpC's share of

valenced speech is 28.7%). A two-sided permutation test that randomly reassigns valence labels across the pool produces no draw with VpC Negative count as high as 122 across all 10,000 simulations; the implied two-sided $p < 10^{-4}$, and the analytic one-sided hypergeometric $p < 10^{-50}$. The bootstrap 95% CI on VpC's Negative rate among VpC valenced invocations is $[87.9\%, 96.2\%]$; the corresponding CI for non-VpC is $[12.5\%, 20.7\%]$. The two intervals do not overlap.

**Cuba—Acosta's share of Negatives.** The 123-invocation valenced pool contains 22 Negatives, 14 from Acosta. Acosta produces 50.4% of valenced invocations; under the null of random label assignment his expected Negative count is 11.09. The two-sided permutation $p$-value on his Negative count is 0.242, and the one-sided hypergeometric $p = 0.128$. The bootstrap 95% CI on Acosta's share of all Negatives (63.6% point estimate) is $[45.5\%, 81.8\%]$, narrowly containing 50%. Conditional on his volume of valenced speech, his Negative count is not statistically distinguishable from chance—the substantive interpretation in the main text is that his dominance of the Negative column reflects floor-time concentration, not an elevated per-invocation Negative rate.

**Cuba—Drafting Commission vs. non-DC Negative rate.** Fisher's exact on the 123-invocation contingency yields OR $= 1.043$, $p = 1.000$ (DC: 18 of 100 valenced; non-DC: 4 of 23). Excluding Acosta from the DC cell yields OR $= 0.559$, $p = 0.461$ (DC: 4 of 38; non-DC: 4 of 23 unchanged). The DC–non-DC Negative-rate difference is statistically and substantively indistinguishable from zero with or without Acosta in the sample.

### A.6.2 Robustness to borderline reclassifications (Cuba)

Four Cuba SD invocations are borderline cases reclassified from Positive to Negative in human review, following the codebook's rule that valence anchors to the speaker's orientation toward the provision being argued rather than toward the consultation input itself (§A.2.5). Each speaker embraced the cited input but proposed an amendment to the article under debate (Caballero Garzón on Article 71, González Llort on Article 30, Saladrigas González on Article 55, counted as two invocations). These differ from outright rejections of proposed provisions and are flagged with the audit-trail `source` field in the analytic CSV.

The pre-override valenced pool contains 18 Negatives instead of 22, plus one additional SD/Positive: the Adán Roble invocation excluded from the analytic dataset as a historical-process reference per §A.2.3 but restored here for the comparison. Acosta's Negative count is unchanged at 14—none of his invocations were affected by the overrides—so his share of all Negatives rises from 63.6% post-override to 77.8% pre-override (14/18). Conditional on his 50% share of the pre-override valenced pool ($K = 62$, $N = 124$), the two-sided permutation test on his Negative count returns $p \approx 0.020$ (one-sided hypergeometric $p \approx 0.010$), against the post-override $p = 0.242$ reported in the main text. The two codings yield different conclusions at the conventional 5% threshold.

The post-override coding is the codebook-correct read, and the four reclassified cases are themselves substantively important: they are non-Acosta floor deputies using consulta input as warrant for amendments to the draft—exactly the kind of dissent that the provision-anchored valence rule is designed to register. Counting them as Negative dilutes Acosta-only concentration of dissent and is what produces the main-text claim that his concentration is consistent with chance. The pre-override result is reported here transparently because it shows the magnitude of the swing the codebook's valence-anchor revision (§A.2.10) drives in this corpus, and it makes the pre-override input-anchored reading legible to readers who would prefer it.

### A.6.3 Paraphrased-prompt replication

The paraphrased-prompt run (Run 4 in Table 1) was executed on the same 13-session Chile subsample and full Cuba corpus used for self-consistency runs. Its agreement with canonical runs was reported in §A.4.4: classification-level $\alpha$ exceeds 0.91 in both corpora, with the only level-2 (count) drift attributable to occasional differences in segmentation between the two prompts. Headline patterns hold under the paraphrased prompt: VpC produces 92%+ of paraphrased-run Chile Negatives in the subsample, and the Cuba paraphrased-run Negative count is dominated by Acosta in the same proportion as the canonical runs. No headline statistical claim flips sign or significance under the paraphrased prompt.

### A.6.4 Specification robustness on the Chile valence logit

The invocation-level Chile valence logit (Table 7) is reported in three specifications because continuous ideology and the binary VpC indicator are highly collinear—VpC delegates sit on the right of the ideology scale—so the joint-coefficient read in Model 3 is hard to interpret in isolation. Pseudo-$R^2$ rises from 0.379 (ideology only) to 0.433 (VpC only) to 0.443 (both). The gain from adding VpC to the ideology-only model (+0.064) is approximately six times the gain from adding ideology to the VpC-only model (+0.010). The substantive read in the main text—that cherry-picking is organized around bloc membership specifically rather than as a smooth function of left-right position—rests on this asymmetric model-comparison read, not on the joint coefficients.

## A.7 Logistic regressions

This section reproduces the full logistic-regression specifications referenced in the main text. The Chile model (Table 7) tests whether bloc membership predicts the valence of an invocation conditional on speaker covariates; the Cuba model (Table 8) tests whether elite status predicts the per-turn propensity to invoke at all. Both estimations cluster standard errors at the speaker level.

### A.7.1 Chile—invocation-level valence logit

The unit of analysis is the valenced invocation ($n = 460$). The outcome is binary: Negative ($= 1$) vs. Positive ($= 0$). Table 7 reports three specifications. Model 1 enters continuous ideology only; Model 2 replaces it with a binary Vamos por Chile indicator; Model 3 includes both. All specifications include the same control set: leadership status (Mesa or commission coordinator), participation-commission membership, and the log word count of the enclosing speaker turn. Standard errors cluster on speaker ID (116 clusters, the number of distinct delegates producing valenced invocations). The marginal model-comparison read is the principal interpretive frame; see §A.6.4.

### A.7.2 Cuba—turn-level invocation propensity logit

The unit of analysis is the deputy speaker turn ($n = 261$; 4 non-deputy turns dropped). The outcome is `has\_invocation` ($= 1$ if the turn produced at least one invocation). Table 8 reports

Table 7: Chile: invocation-level logistic regression of Negative (= 1) vs. Positive (= 0) valence.

| | Model 1<br>*ideology only* | Model 2<br>*VpC only* | Model 3<br>*both* |
|---|---|---|---|
| *Predictors of interest* | | | |
| Ideology (continuous) | 14.689*** | — | 2.840* |
| | (7.030, 30.694) | | (1.025, 7.870) |
| Vamos por Chile | — | 47.914*** | 18.063*** |
| | | (17.858, 128.557) | (5.003, 65.212) |
| *Controls* | | | |
| Leadership role | 0.496† | 0.696 | 0.727 |
| | (0.242, 1.015) | (0.369, 1.313) | (0.377, 1.405) |
| Participation commission | 0.547 | 0.621 | 0.614 |
| | (0.214, 1.399) | (0.251, 1.534) | (0.231, 1.636) |
| Log word count | 0.401*** | 0.462*** | 0.446*** |
| | (0.260, 0.618) | (0.294, 0.727) | (0.287, 0.693) |
| Constant | 297.724*** | 27.888* | 52.287** |
| | (19.371, 4575.911) | (1.871, 415.770) | (3.533, 773.827) |
| Invocations ($N$) | 460 | 460 | 460 |
| Negatives | 176 | 176 | 176 |
| Speaker clusters | 116 | 116 | 116 |
| Pseudo-$R^2$ | 0.379 | 0.433 | 0.443 |

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Notes:* Cells report odds ratios with 95% cluster-robust confidence intervals in parentheses (clustering on speaker ID, 116 clusters). Estimation uses the 460 valenced invocations with non-missing `ideology1`, `ID`, and `word\_count`; 176 are Negative. Model 1 enters continuous ideology only; Model 2 replaces it with a binary Vamos por Chile indicator; Model 3 includes both.

two specifications. Model A enters Drafting Commission and Central Committee membership as dichotomous elite-status indicators alongside log turn duration. Model B orthogonalizes Acosta from the remaining DC and CC effects by replacing the two indicators with their non-Acosta counterparts and adding an Acosta-specific dummy. Standard errors cluster on speaker ID (70 clusters). The Acosta odds ratio of 0.347 ($p = 0.019$) in Model B means that, conditional on turn duration, his per-turn odds of producing an invocation are roughly one-third of a baseline non-elite speaker—consistent with the interpretation that his dominance of the Negative column is driven by speech volume rather than per-turn invocation density.

Table 8: Cuba: turn-level logistic regression of invocation propensity (`has_invocation` = 1 if the turn produced ≥1 invocation).

| | Model A<br>*core* | Model B<br>*Acosta-specific* |
|---|---|---|
| *Predictors of interest* | | |
| Drafting Commission | 1.312 | — |
| | (0.442, 3.899) | |
| Central Committee | 0.799 | — |
| | (0.191, 3.343) | |
| Drafting Commission (excl. Acosta) | — | 1.756 |
| | | (0.560, 5.509) |
| Central Committee (excl. Acosta) | — | 2.422 |
| | | (0.742, 7.905) |
| Acosta | — | 0.347* |
| | | (0.144, 0.837) |
| *Controls* | | |
| Log turn duration | 5.026*** | 6.733*** |
| | (3.458, 7.304) | (3.509, 12.918) |
| Constant | 0.000*** | 0.000*** |
| | (0.000, 0.001) | (0.000, 0.001) |
| Turns ($N$) | 261 | 261 |
| Speaker clusters | 70 | 70 |
| Pseudo-$R^2$ | 0.455 | 0.497 |

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Notes:* Cells report odds ratios with 95% cluster-robust confidence intervals in parentheses (clustering on speaker ID, 70 clusters). Outcome is `has_invocation`: 1 if the turn produced ≥1 invocation. Estimation uses 261 deputy speaker turns (4 non-deputy turns dropped). Model A enters Drafting Commission and Central Committee membership as dichotomous elite-status indicators alongside log turn duration; Model B orthogonalizes Acosta from the remaining DC and CC effects by replacing the two indicators with their non-Acosta counterparts and adding an Acosta-specific dummy.

## A.8 Spanish-language originals of block quotes

The block quotes appearing in English in the main text were translated by the author from Spanish-language plenary transcripts (Chile) and ASR-transcribed Facebook livestreams (Cuba). The originals are reproduced below in the order they appear in the main text. Citations identify the speaker, session date or session number, and the row in the source transcript.

### A.8.1 Chile

He puesto mucha energía en realizar cabildos con jóvenes del distrito 26, muchos de ellos no votarán en este plebiscito de salida, pero sí convivirán con el trabajo que florezca desde esta Convención. En todos estos cabildos mencionaron la necesidad de la educación sexual integral. Una joven en Quinchao me dijo: Adriana, si no se escribe una Constitución que incluya la educación sexual integral, entonces todas las jóvenes del país corremos grave riesgo, pues, los compañeros deben aprender sobre consentimiento y respeto. Hasta hoy sus palabras acompañan mis reflexiones. Me fui a Calbuco, misma preocupación de los jóvenes del Colegio Goycolea, educación sexual integral como demanda prioritaria. (Adriana Ampuero, Sesión 68 Plenario, row 143)

En estos momentos quiero llamar a aprobar la norma proveniente de la iniciativa popular sobre el derecho a la vivienda, la función social de la propiedad del suelo y la producción social del hábitat. … Tuvo 22.000 firmas de apoyo, gestada a través de cabildos, de asambleas y de organización popular. Este pueblo organizado que presenta esta propuesta y sostiene esta demanda es el pueblo movilizado y organizado que luchará por el *Apruebo* de salida. (Manuela Royo, Sesión 85 Plenario, row 211)

Yo le pregunto ¿y a cuántos de ese millón escucharon en esta Convención si rechazaron muchas de las normas con mayores apoyos? ¿Dónde está la opinión mayoritaria de la ciudadanía aquí reflejada? Les dijeron que no a "Con mi plata no," iniciativa popular que con más de 60.000 apoyos fue presentada; la iniciativa por el libre derecho de la propiedad privada, con más de 47.000 apoyos; le dijeron que no a "Primero las víctimas," con más de 26.000 apoyos; Libertad de Emprendimiento, la Multigremial de Emprendedores; Educación Libre y Diversa; y un gran, gran, gran, etcétera. Una ficción de escuchar a la ciudadanía, pero algo más que ficción no le podíamos pedir a un escritor de ficción. (Katerine Montealegre, Sesión 78 Plenario, row 136)

Todo este proceso de elaboración de normas estuvo abierto a la participación ciudadana, desde las audiencias que recibieron las distintas comisiones hasta el proceso de consulta indígena y la presentación de iniciativas populares de norma. ... Luego del análisis de los datos de deliberación y votaciones que se efectuaron en las distintas comisiones temáticas ... podemos señalar que un 49 por ciento de las iniciativas populares de norma fueron aprobadas de forma total o parcial. De las restantes 42 iniciativas, fueron rechazadas en general en el trámite de las comisiones y una declarada inadmisible. Al analizar su contenido, es posible determinar que el 83 por ciento de ellas contenían propuestas, ideas y normas que fueron aprobadas por el Pleno de la Convención Constitucional y, por tanto, ya integran el borrador de la nueva Constitución. De esta forma, el 91,5 por ciento de las propuestas populares que han sido discutidas por este órgano constituyente logró incidir de manera efectiva en el debate de las normas. (María Elisa Quinteros, Sesión 96 Plenario, row 2)

Ya se ofició sobre los cabildos y encuentros locales autoconvocados que no tuvieron incidencia. Esta misma inclusive intentó aumentar el número de participantes, que no superó los tres mil, sumando incluso el que se hizo en el gobierno anterior. Lamentablemente a esto se suma la jornada nacional de deliberación que no cumple con los principios que esta Convención se propuso. Esta jornada no es incidente ... y dudo que sea eficiente. ¿Qué incidencia tiene esta jornada en un texto que ya tiene cerca de 400 artículos y que está prácticamente listo? ¿Es oportuno hacerlo al final del proceso? Evidentemente, no. (Felipe Mena, Sesión 97 Plenario, row 12)

Ya lo hemos advertido antes y se los hemos repetido hasta el cansancio: para nosotros este informe no tiene ninguna validez, porque proviene de una consulta indígena mal hecha, en la que participó menos del uno por cierto del padrón. Por eso, como bancada hemos decidido no votarlo, porque no estamos dispuestos a validar un proceso viciado como este. (Eduardo Cretton, Sesión 101 Plenario, row 95)

### A.8.2 Cuba

De forma general la población, todo nuestro pueblo ha reiterado en disímiles ocasiones, lo ha hecho en los procesos de consulta y también en distintos escenarios ha reiterado y ha elogiado el carácter democrático de este proceso de consulta popular. Le decía la singularidad que tiene este proceso, que nosotros lo decimos sin vanidad, lo decimos con humildad, pero tenemos que reiterarlo. Es un proceso singular que no existe ni se da en ningún lugar de este mundo. En que se convoque a un pueblo, a escuchar a un pueblo, a oír la opinión de un pueblo sobre un proyecto constitucional. Y no solo que se escuche, sino que se tome en cuenta la opinión del pueblo. (Homero Acosta Álvarez, 21 December 2018, row 2)

De la manera en que se consultó, nunca se pidió a las personas que dieran aprobación sobre lo que estaba en el texto. Por tanto, siempre hay una zona de duda, una zona que no hay una certidumbre absoluta de hasta dónde se puede apoyar o no un artículo en concreto. Porque no se votaba los que estaban a favor o no de un artículo u otro artículo. Esa fue la manera en que se procesó. Y por eso es que aquí solamente salen las personas que opinaron contrario a ese artículo. También hay muchas personas que tenían una opinión diferente o que también tenían esa opinión similar y tampoco la expresaron … Por tanto, ahí ya le digo, hay siempre una zona en que queda cierto empirismo en hasta dónde llega esa opinión en el cúmulo de la población que participó dentro de la consulta popular. (Homero Acosta Álvarez, 21 December 2018, row 3)

Bueno, yo creo que en definitiva hemos ganado mucho y agradecemos a todas las personas que en la consulta popular expresaron sus opiniones, ese 24%. Yo creo que también tenemos que agradecerle que con transparencia en la consulta hayan planteado sus opiniones porque ese es el reflejo de su sentimiento, de su sentir independientemente de que quizás la mayoría no coincidamos, pero eso también … En la vida es así, no tenemos por qué estar todas de acuerdo y todos con todo lo que se propone. Yo

creo que fue un ejercicio democrático importante el que hemos hecho y hemos ganado. ¿Por qué hemos ganado? Primero porque hemos expulsado el sentir de la población y sabemos cómo se piensa para poder ajustar todas las políticas públicas que hay que ajustar y sobre todo todo el trabajo de sensibilidad. (Teresa María Amarelle Boué, 20 December 2018, row 50)

Pero en función de eso y en la medida que hacíamos nosotros nuestro proceso de consulta popular, lo publicaba el *Nuevo Herald*, lo publicaba sitios como Cubanet, se realizaban también en paralelo sondeos dirigidos y a través de aplicaciones en los móviles a cubanas y cubanos que manifestaban su desacuerdo con … Por ejemplo, tener un partido único con el carácter irrevocable del socialismo o con que la principal forma de propiedad en nuestro país fuera la socialista. Lamentablemente para ellos, bueno, tenían que reconocer que ese sondeo solamente había alcanzado alrededor de 1.612 personas, aunque sus titulares manifiestan que era un 47% del pueblo cubano el que estaba en desacuerdo con lo que estábamos nosotros analizando, con lo que estábamos planteando y refrendando en la propia Constitución. Nada que ver con las cifras que magistralmente ayer usted mostraba y que apuntan, además del 62% este de criterios favorables, cuán en sintonía y cuán de acuerdo estamos con lo que estamos construyendo. (Reina de la Caridad Torres Pérez, 22 December 2018, row 13)

Bueno, aquí sería bueno comentar algo, teniendo en cuenta la cantidad de opiniones que se recibieron durante el ejercicio de la consulta popular, que estaban asociadas a varias propuestas, y no de una única manera, sino de diferentes formas también, de plantear la elección directa del presidente de la República. … Sobre estos planteamientos de ser más democráticos o menos democráticos en el ejercicio de la elección directa del presidente, miren, tenemos que sentirnos orgullosos todos, hombres y mujeres en Cuba, de tener una experiencia democrática muy amplia en estos 60 años de revolución. Y tenemos también la historia vivida de comparación con el ejercicio antes

de 1959 para saber qué cosas eran las elecciones en Cuba, con eso que tanto nos propugnan a veces que tengamos, que es los partidos electorales. Esa experiencia ya Cuba la vivió. Viviendo ahora y refrescando y recreando y perfeccionando, como el comandante en jefe siempre nos dijo y nos dice … hay que pensar que para ser democrático lo que hay que tener es un sistema político, económico, social y también cultural que defienda los intereses de todas las personas. (Ana María Mari Machado, 20 December 2018, row 61)

Nos vamos a referir en nuestra intervención al artículo 55 referido a la libertad de prensa, a partir no sólo de criterios personales, sino sobre todo a lo que escuchamos en el proceso de consulta popular. … Tenemos la percepción de que fue uno de los reclamos fundamentales que desde el gremio, desde la Unión de Periodistas de Cuba, se realizó durante el proceso de consulta. O sea, que la Constitución dejara bien claro que los medios fundamentales de comunicación en Cuba solamente podrían ser del pueblo. Por lo tanto, nos parece más conveniente que quede redactado el final de ese artículo de la siguiente manera: no pueden ser objeto, los medios fundamentales, como dice el artículo, no pueden ser objeto de ningún otro tipo de propiedad. (Daicar Saladrigas González, 22 December 2018, row 106)

## A.9 Data and code availability

The replication package accompanying this article includes:

- The full codebook (v2.7) reproduced in §A.2, plus the changelog of all prior versions (§A.2.10).
- The canonical and paraphrased system prompts (§A.3.1, §A.3.3) and country-specific guidance (§A.3.2).
- The pipeline source (`scripts/pipeline/code\_transcripts.py`), validation scripts (`scripts/validation/`), and analysis scripts (`scripts/analysis/`) used to generate every result reported in the main text and this appendix.

- The raw API outputs (one JSON file per session, in `data/output/\{country\}/run\_\{1--4\}/raw/`), the flat CSV (`data/output/\{country\}/run\_\{1--4\}/csv/`), and the analytic dataset with merged speaker metadata (`analysis/\{country\}/\{country\}\_analytic.csv`).

- The processing logs for every run (`data/output/logs/`), recording status, invocation count, JSON-repair flags, unlocatable-excerpt counts, and timestamps for every speaker turn processed.

- The validation artifacts: stratified samples (`data/validation/\{country\}/sample\_\{country\}.csv`), hand-coded reviews (`review\_\{country\}.csv`), human-overrides files (`human\_overrides\_\{country\}.csv`), the lexical recall floor outputs (`lexical\_recall\_floor\_\{country\}.csv`), and the cross-run comparison data (`\{country\}\_run\_comparison.csv`).

- The Cuba transcripts produced by WhisperX from the *Cubadebate* broadcasts. Source audio is publicly archived but not redistributed; the transcripts are released as the analytic record.

- The Chile transcripts as obtained from the Library of the National Congress of Chile, retained as the analytic record. Re-distribution rights are governed by the original archive's terms.

The full pipeline is reproducible from the released materials with fixed seeds (`20260421` for Chile, `20260420` for Cuba), 10,000 permutation / bootstrap draws, and a pinned codebook version (v2.7) and prompt variant (canonical for Run 1; paraphrased for Run 4). The release URL and DOI will be added to the published version of this article.

# References

Bain, Max, Jaesung Huh, Tengda Han, and Andrew Zisserman. 2023. *WhisperX: Time-Accurate Speech Transcription of Long-Form Audio.* ArXiv:2303.00747 [cs, eess]. Accessed May 25. http://arxiv.org/abs/2303.00747.

Fábrega, Jorge. 2022. *Sesgos, falsedades y prejuicios ante el trabajo constituyente: así comentan las redes.* Section: Columna. Accessed September 25. https://www.ciperchile.cl/2022/04/22/sesgos-falsedades-y-prejuicios-ante-el-trabajo-constituyente/.

Krippendorff, Klaus. 2019. "Reliability." In *Content Analysis: An Introduction to Its Methodology,* 4th ed. SAGE Publications, Inc. Accessed April 30.

Radford, Alec, Jong Wook Kim, Tao Xu, Greg Brockman, Christine Mcleavey, and Ilya Sutskever. 2023. "Robust Speech Recognition via Large-Scale Weak Supervision." In *Proceedings of the 40th International Conference on Machine Learning,* 28492–28518. PMLR. Accessed October 1. https://proceedings.mlr.press/v202/radford23a.html.

Rozas, Joaquin, Alejandro Olivares L., and Antoine Maillet. 2023. "Entre la independencia y la militancia: análisis de la vinculación partidaria de los convencionales constituyentes de Chile." *Economía y Política* 9 (2): 67–100.