

FrIdéo: French News Outlet Ideology Scoring via Multi-Source Evidence Fusion and Distancing-Aware Lexical Scoring

Amr Sobhy

Le French News Lab
amr@frenchnews1ab.org

Abstract

Outlet ideology is a standard explanatory variable in media framing, media effects, and political communication research, but continuous, ecosystem-wide coverage lets it enter as graded exposure rather than a coarse bin. For France, existing references are either broad but non-analytic or analytic but sparse and source-specific, and none we are aware of offers ecosystem-wide coverage on a unified continuous scale. We contribute FrIdéo: to our knowledge the first ecosystem-wide continuous ideology scores for the French press, a transparent composite score for 30 French news outlets, fusing nine independent families of evidence with empirical-Bayes-style shrinkage toward a structural prior. For the text-based family we introduce *symmetric plain/qualified scoring*, which separates the plain use of ideologically loaded vocabulary from its qualified (distanced) use on both sides, a distinction bag-of-words methods miss. The resulting scale runs from L’Humanité on the far left to Valeurs actuelles on the far right; the far-right pole is near-indistinguishable and CNews ties Le Figaro. Because the text family reads current articles, it also captures editorial change that static sources miss, such as the *Journal du Dimanche*’s sharp rightward shift after its 2023 ownership change. The ordering is robust to perturbation and to dropping any single family, and agrees with held-out Media Bias Fact Check labels. We release the scores, confidence intervals, and replication code.

1 Introduction

The political orientation of news outlets is among the most consequential variables in political communication research. It shapes event selection, framing, source amplification, and audience belief updating. Continuous, ecosystem-wide scores let ideology enter as a graded exposure rather than a coarse bin (Martin and Yurukoglu, 2017),

and avoid selecting on whichever outlets a rating panel happens to cover. Yet scoring outlet ideology is methodologically fraught: outlets resist self-categorization, audience surveys conflate perceived vs. actual bias, and content analyses require large corpora and fine-grained annotation.

The dominant paradigms in the English-language literature fall into four camps. *Audience-based methods* infer outlet ideology from the partisanship of its readers or social-media followers (Bakshy et al., 2015). *Content-based methods* analyze word choice, topic emphasis, or citation patterns, or crowd-code article content, to place outlets on a liberal–conservative axis (Groseclose and Milyo, 2005; Gentzkow and Shapiro, 2010; Budak et al., 2016). *Network-based methods* derive positions from citation, follower, or web-link graphs linking outlets to political actors (Cointet et al., 2021; Fondation Descartes, 2020). *External-rating methods* rely on crowd-sourced or expert panels such as Media Bias Fact Check (MBFC) or AllSides (Media Bias Fact Check, 2026). Each has known failure modes: audience methods conflate readership ideology with editorial stance; content methods require large corpora and struggle with outlets that report without endorsing; network methods inherit the snapshot date of the graph they are built on; rating panels are U.S.-centric and sparsely cover European media.

For the French media ecosystem, the situation is acute. France has dozens of nationally relevant news outlets spanning print, broadcast, and digital-native formats, covering the full spectrum from the Communist *L’Humanité* to the identitarian *Valeurs actuelles*. Available references for outlet ideology in France are either broad but non-analytic (descriptive typologies without a measurement model) or analytic but sparse and source-specific: Pew Research Center (2018) place only eight outlets on a perceived-ideology scale; Cagé et al. (2024) estimate partisan deviation for broad-

cast outlets only, from TV and radio guest bookings; Cointet et al. (2021) derive continuous positions via network analysis but from a static 2018–2019 snapshot; Benson and Hallin (2007) offer a comparative content analysis of only three outlets. No single source currently combines broad ecosystem coverage with a unified continuous scale for the sampled French national news ecosystem.

This paper contributes a resource: FrIdéo, to our knowledge the first continuous, ecosystem-wide ideology dataset for French national news, a latent, anchor-free, z-standardized score with a per-outlet 95% bootstrap confidence interval (CI) for 30 outlets, together with the method used to build it. It is validated through cross-family convergence, known-groups separation, and robustness checks; held-out MBFC agreement is secondary external corroboration, not the target being reconstructed. Two methodological components make it possible: a unified multi-source framework that fuses nine evidence families into a single latent score via per-family z-standardization and empirical-Bayes shrinkage toward a structural prior; and a signal-based lexical score (D7) that runs on raw article text and distinguishes *plain use* of ideologically loaded vocabulary (no explicit distancing signal present) from *qualified use* (guillemets, distancing-prefix, or side-specific critique vocabulary present), yielding a time-sensitive editorial signal that registers within-period editorial shifts (e.g., JDD’s 2023 realignment) unrecoverable from static structural priors or historical typologies; validated against a manually adjudicated gold standard it reaches an operational-fidelity $\kappa = 0.938$ (0.78 on human-adjudicated rows alone), exceeding both an LLM judge ($\kappa = 0.49$) and a CamemBERT probe (macro-F1 = 0.775) on this narrow signal-detection task. We follow the outlet-scoring literature in treating ideology as a single left–right continuum. This is an explicit modeling choice, which we test rather than assume (Section 6, Appendix F.3), and which can collapse distinct economic and cultural axes where these diverge. We conjecture the distancing check transfers to other languages with quotation conventions and a lexicalized political register, but D7 is validated on French items only, so that transfer is untested. Unlike phrase-co-occurrence and citation-pattern approaches that anchor to a reference corpus of ideologically tagged speech (Gentzkow and Shapiro, 2010; Groseclose and

Milyo, 2005), FrIdéo imposes no external anchor. FrIdéo is a resource-building step, not a definitive settlement.

2 Related Work

Groseclose and Milyo (2005) and Gentzkow and Shapiro (2010) estimate U.S. outlet slant from citation patterns and phrase co-occurrence with congressional speech, but both require large ideological reference corpora that do not transfer to French media. Budak et al. (2016) crowd-code article content; Rönnback et al. (2025) train a classifier on outlet-level GDELT features with MBFC ground-truth labels. FrIdéo differs by treating outlet ideology as a *composite latent position* derived from nine heterogeneous evidence families rather than optimizing a single content model, making each source’s directional contribution transparent and accommodating sparse content coverage, at the cost of multi-source evidence assembly. Event-level framing work (Zhao et al., 2025; Das et al., 2024; Concha Macías and Norambuena, 2024) addresses the distinct problem of how a given event is framed across outlets, complementing our outlet-level scores.

Editorial change over time. Most outlet-scoring work treats ideology as static within the estimation window, but a smaller literature does not: Martin and Yurukoglu (2017) recover MSNBC’s post-2006 leftward move from time-varying slant estimates, and Kim et al. (2022) show slant moves substantially even over short horizons. That outlets shift is therefore established; our contribution is integrating it into a multi-source model, pairing one within-period text family with eight static ones so an ownership change registers in the fused score instead of being averaged away. For French we are aware of no existing time-varying measure.

For the French context, Benson and Hallin (2007), working within the comparative-media-systems tradition, characterize the French press as more closely tied to the political field than its US counterpart. Cagé et al. (2022, 2024) provide among the most rigorous empirical evidence for French outlet ideology, using panel data on TV and radio guest bookings to estimate partisan deviation relative to a public-broadcaster benchmark. Cointet et al. (2021) apply correspondence analysis to the Twitter-follower network of 831 MPs to place outlets on an economic left-

right axis. [Pew Research Center \(2018\)](#) provides survey-based perceived-ideology scores, and [Fondation Descartes \(2020\)](#) maps the French politico-media space from Twitter follower-interaction networks. [Rebillard and Sklower \(2022\)](#) and [Cagé \(2017\)](#) document ownership concentration and its political implications. Our framework synthesizes all of these into a common scale.

Lexical content-scoring methods often confuse *reporting about* partisan vocabulary with *plain use* of it (for a review of bias-detection methods, see [Hamborg et al., 2019](#)), since bias is often sentence-level ([Lim et al., 2020](#)): an outlet that covers far-right demonstrations must mention *ensauvagement* (‘descent into savagery’) or *grand remplacement* (‘great replacement’) to describe them; counting these as plain use inflates its right score. Our plain/qualified procedure (Section 5) applies a three-layer distancing check before crediting an occurrence. We adapt James–Stein shrinkage ([James and Stein, 1961](#); [Efron and Morris, 1977](#)), in the spirit of latent ideal-point estimation in political science ([Jackman, 2001](#); [Shor and McCarty, 2011](#)), to heterogeneous evidence with variable coverage: rich-evidence outlets weight toward the empirical mean, sparse ones toward a historical-tradition prior.

3 The French Outlet Panel

Our panel consists of 30 French-language national news outlets: 15 national newspapers and magazines, 6 broadcast TV/radio news channels, 8 digital-native outlets, and 1 cooperative regional paper (see Table C.1 for the full list). The selection aimed for broad coverage of the national news ecosystem rather than a threshold-filtered sample. Inclusion criteria were: national or near-national reach; a general political or current-affairs mandate (excluding sports, lifestyle, and trade outlets); representation of all major formats (daily and weekly press, broadcast, digital-native, regional press); and coverage of the full ideological spectrum from far-left (*L’Humanité*) to identitarian far-right (*Fdesouche*, *Valeurs actuelles*). No minimum audience threshold was applied. No stratum is an exhaustive enumeration of its format: coverage of the ACPM-ranked national dailies is complete (below), while the magazine, broadcast (6) and digital-native (8) strata are purposive selections. *Ouest-France* is the single regional title because it outsells every national paper in the

same ACPM ranking; its cooperative ownership is incidental, not the inclusion criterion. The panel includes every title in the ACPM (Alliance pour les Chiffres de la Presse et des Médias) 2025 paid-circulation and 2026 readership rankings¹ except the sports daily *L’Équipe* and the international edition of the *New York Times*, both outside our general-news, French scope; *L’Opinion* is additionally included as a national daily below ACPM’s reporting threshold.

The panel includes outlets under a range of ownership structures: state/public, cooperative/non-profit, legacy independent, and oligarchic (Bolloré/Vivendi, Arnault/LVMH, Dassault, Drahi, Pinault). Ownership is treated as one evidence family (D0) rather than as a model constraint; removing D0 entirely yields leave-one-out (LOO) $\rho = 0.997$, confirming it is not load-bearing.

4 The Ideology Scoring Framework

4.1 Evidence Families

We define nine direct indicator families (with D6 split into two sub-families, D6a and D6b) and one structural prior (S2). Table A.1 summarizes each family.

Source consistency parameters σ_k are estimated via leave-one-family-out residual variance: $\sigma_k^2 = \text{Var}(z_{ik} - \bar{z}_{i \setminus k})$ across outlets with data in k , where $\bar{z}_{i \setminus k}$ is the mean of all other families. Lower σ_k means the family tracks the aggregate closely; higher σ_k reflects divergence—noise or genuine signal absent from the consensus. This is a modeling assumption, not a ground-truth criterion. Families capturing signal absent from historical sources—notably D7, the only within-period lexical family—exhibit larger σ_k (D7: 0.42). D1 covers only 3 outlets, so its σ_k is an unstable conservative default; it carries negligible weight but is retained for completeness. D3 is our dual-PageRank score over the 2018–2019 [Cointet et al. \(2021\)](#) web-link network (not Cointet’s own Twitter-follower position), and is therefore stale for outlets that moved after 2019. The corrective is not D7 alone: a family-level decomposition (Appendix E) shows D7 contributes most at JDD but only fourth-most at CNews, where D0 and D6b carry more of the mass. Owner-

¹ACPM *Diffusion Presse Quotidienne Nationale* 2025 and *OneNext PQN* 2026-S1, <https://www.acpm.fr/c1assements>.

ship (D0) is mapped through a discrete owner-structural lookup table (20 scored categories plus two null categories—for concealed ownership or an ideologically-unmarked owner, which receive null D0; the complete owner-category mapping is in the replication package). Because D0 is z-standardized, only the ordinal ranking of owner types matters (Communist Party (PCF) organ < independent left < public/cooperative < conservative oligarch < far-right aligned); see Table B.1 for per-family LOO details.

4.2 Per-Family Z-Standardization

Each evidence family k is standardized independently across the subset of outlets with available data:

$$z_{ik} = \frac{x_{ik} - \bar{x}_k}{s_k}$$

where $\bar{x}_k$ and s_k are the population mean and standard deviation computed over the n_k outlets with non-null data in family k . Families with fewer than two observed values are set to zero (uninformative). Outlets missing family k are absent from that family's z-matrix entirely; they do not receive an imputed zero, which would introduce spurious shrinkage toward the center. For some families missingness is structural (D1 covers only specific outlet types), not random; since scores are never imputed, no missing-at-random assumption is needed.

This ensures comparability across heterogeneous scales (D2: 0–6 survey; D1: percentage points; D7: ± 1 ratio); per-family z-standardization is canonical in composite indicator construction (OECD and European Commission Joint Research Centre, 2008).

4.3 Latent Score via Empirical-Bayes Shrinkage

FrIdéo is a composite-indicator measurement model that estimates a latent ideological position from multiple heterogeneous evidence families; unlike factor-analytic or ideal-point models, it does not infer positions through a generative statistical model.

For each outlet i , a blended latent score weights direct evidence against a historical editorial tradition prior: the more evidence families cover an outlet, the closer its score tracks their mean:

$$\begin{aligned}\bar{z}_i^{\text{direct}} &= \frac{1}{n_i} \sum_{k \text{ avail.}} z_{ik} \\ z_i^{\text{struct}} &= 1.5 S2_i, \quad \alpha_i = \frac{n_i}{n_i + \lambda} \quad (\lambda = 2) \\ z_i^{\text{blended}} &= \alpha_i \bar{z}_i^{\text{direct}} + (1 - \alpha_i) z_i^{\text{struct}}\end{aligned}$$

The shrinkage parameter $\lambda = 2$ implies that an outlet with two direct families receives equal weight between empirical evidence and the structural prior ($\alpha = 0.50$); four families yields $\alpha = 0.67$; seven families yields $\alpha = 0.78$. Outlets with no direct evidence ($n = 0$) fall back entirely on the structural prior. The structural prior S2 encodes founding ideological tradition on a symmetric five-point ordinal grid (-0.5 to $+0.5$), scaled by 1.5 to match the empirical z-score spread. S2 values were assigned by the authors from each outlet's documented founding profile—founding year, founder or institution, initial editorial alignment—using published press histories and editorial archives; the full per-outlet coding is in Appendix G.

The prior acts as a regularizer pulling low-evidence outlets toward their documented historical identity, while high-evidence outlets are determined primarily by the direct pool. $\lambda = 2$ is a design choice, not an estimate: it sets $\alpha = 0.50$ at the minimum meaningful evidence count ($n_i = 2$), encoding the judgment that two sources should not override the historical record without corroboration. We use *empirical-Bayes* in the estimation sense that the family dispersions σ_k are data-estimated (leave-one-family-out residual variance), while λ and the prior mean (S2) are fixed by design; the procedure is thus *empirical-Bayes-style*, not a fully estimated hierarchical model. Shrinkage depends on evidence quantity (n_i), not inter-family coherence, keeping point estimation and uncertainty quantification separate. $\bar{z}_i^{\text{direct}}$ is *unweighted* rather than precision-weighted ($1/\sigma_k^2$): since σ_k is largest for families that diverge from consensus (notably D7), precision weighting would suppress the contemporary signal the model is designed to keep auditable.

The blended scores are re-standardized across all 30 outlets:

$$\text{final}_i = \frac{z_i^{\text{blended}} - \mu_B}{\sigma_B}$$

where μ_B and σ_B are the mean and population standard deviation of $\{z_i^{\text{blended}}\}$. The result is

a zero-mean, unit-variance scale across the 30-outlet panel. No external anchors constrain this scale; the outlets at its extremes (L’Humanité on the left, Valeurs actuelles on the right) are validated post-hoc, not used to fix it.

The scale is panel-relative. Because every family is z-standardized over the panel and the blended scores re-standardized across the same 30 outlets, the zero-point and unit are defined *internally*; an anchor-free scale cannot also be addition-invariant. Adding or removing outlets therefore shifts the reference distribution and can move boundary-proximate outlets across σ -band thresholds even when rank-order is unchanged. FrIdéo scores are thus comparable only *within* a fixed panel: applying the method to a different outlet set re-defines the distribution and requires re-estimation rather than transplanting absolute scores. Externally anchored calibration (Limitations) is left to future work.

Each outlet receives an analytic uncertainty estimate that grows with evidence sparsity:

$$\begin{aligned}\text{var}_i &= \sum_{k \text{ avail.}} \sigma_k^2 \\ \sigma_i^{\text{base}} &= \frac{1}{\sigma_B} \cdot \frac{\sqrt{\text{var}_i}}{n_i} \\ \text{penalty}_i &= \frac{\lambda}{n_i + \lambda} \\ \sigma_i &= \sigma_i^{\text{base}} \cdot (1 + \text{penalty}_i)\end{aligned}$$

The sparsity penalty (penalty_i) ensures that outlets relying heavily on the structural prior are assigned wider confidence intervals. The 95% CI is $[\text{final}_i - 1.96\sigma_i, \text{final}_i + 1.96\sigma_i]$.

Covariance sensitivity. The additive CI formula assumes family independence ($\text{Var}(\bar{z}_i) = n_i^{-2} \sum_k \sigma_k^2$). The three families most likely to covary are the externally-assessed families D2 (audience survey), D4 (bias rating), and D6b (third-party profiling); under a conservative pairwise $\rho = 0.3$ among them, CI half-widths inflate by 3.1% on average (max 12.1%, at Libération, Le Monde and Le Figaro) and no outlet’s label changes. Because the families are by design convergent (mean inter-family $r = 0.84$, Section 6), the independence-based CIs should be read as lower bounds on uncertainty.

Seven ideology labels (*far-left* through *far-right*) are assigned using symmetric σ -band cuts at $\pm 1.5\sigma$, $\pm 0.75\sigma$, $\pm 0.25\sigma$ on the z-standardized

scale, corresponding to the French seven-category political convention (*extrême gauche* ‘far-left’ through *extrême droite* ‘far-right’). The nested widths are chosen so that every category is populated: under the left-skewed French press distribution, equal-width bins would leave the centre category degenerate. All 30 assignments match established characterizations in the literature; a sensitivity check confirms they are stable under alternative symmetric cut-point choices. For the three sparse outlets with a non-zero prior (Blast, Le HuffPost, Atlantico) that agreement is partly circular, the prior encoding the same historical characterization the label is checked against; refitting without the prior leaves their labels unchanged (Appendix E). Slate.fr and Challenges are not circular in this way—their prior is zero, so it exerts no directional pull—but face the converse problem of thin, unanchored evidence, and their labels should be read accordingly. Labels are provided for interpretability only; all substantive analyses use continuous scores. The unidimensional assumption underlying them is supported empirically (principal component analysis, PCA, on the four highest-coverage families; Section 6 and Appendix F.3), though it may collapse distinct economic and cultural dimensions where these diverge (e.g. La Croix and Les Echos, both right-of-center via Catholic tradition vs. economic liberalism; see Limitations).

Because the blended score is a weighted sum, each family’s signed contribution to the final score is exactly recoverable: $C_{ik} = (\alpha_i/n_i) \cdot z_{ik}/\sigma_B$ for direct families, and $C_{i,\text{prior}} = (1 - \alpha_i) \cdot z_i^{\text{struct}}/\sigma_B$ for the structural prior. Appendix D reports these normalized additive contribution shares, grouped into five interpretable evidence classes, for all 30 outlets; absolute shares describe relative evidence weight, signed shares indicate directional pull.

5 Content-Based Lexical Scoring (D7)

D7 is computed from raw article text (headline + body) using regex term matching against 20 left-coded patterns (e.g., *violences policières* ‘police violence’, *fémicide* ‘femicide’, *racisme systémique* ‘systemic racism’, *urgence climatique* ‘climate emergency’) and 15 right-coded patterns (e.g., *ensauvagement* ‘descent into savagery’, *grand remplacement* ‘great replacement’, *islamo-gauchisme* ‘Islamism-leftism’, *immigration massive/incontrôlée* ‘mass/uncontrolled

immigration’). The lexicon’s candidate pool was drawn from a three-month pilot corpus (*politique* ‘politics’-tagged articles, July–September 2023) and curated manually against explicit criteria rather than by any automated correlation search; individual retain/reject decisions were validated against the 2024–2025 corpus and held-out MBFC labels; of roughly 80 candidates, we kept the 35 above and rejected the remainder for neutral saturation (e.g., *délinquance* ‘delinquency’, *émeutiers* ‘rioters’) or polysemy (e.g., *séparatisme* ‘separatism’), with the full rejection log in Appendix H.2.

An occurrence is a *plain use* when its sentence contains none of D7’s distancing signals and a *qualified use* when any fires; absence defaults to plain use. D7 tests only for an explicit distancing signal, not endorsement.

The check applies to every occurrence; sentence boundary extends through any unclosed «...» spanning a period break. Six signals qualify an occurrence: (i) *quotation enclosure*, the term falls inside French guillemets «...», ASCII/curly quotes, an unclosed « within 150 chars, or a colon-plus-quote attribution; (ii) *distancing prefix*, ironic operators such as *théorie du* ‘theory of’, *soi-disant* ‘so-called’, *prétendu* ‘alleged’; (iii) *side-specific critique vocabulary*, e.g. *complotiste* ‘conspiracist’, *réfute* ‘refutes’ for right-coded terms and *supposé* ‘supposed’, *idéologie woke* ‘woke ideology’ for left-coded (full lists in the replication package).

An article-level majority-vote rule applies symmetrically: an article is counted as a left- or right-signal only when plain-use occurrences for that side outnumber qualified-use occurrences.

LLM validation. To validate the distancing classifier, we drew a stratified sample of 527 occurrences from the *politique* corpus (2023-H2 and 2024-H1, distinct from the all-category 2024–2025 scoring corpus), stratified by term side, D7 decision, and outlet ideology band, with outlet identity masked. Each occurrence was judged once by gpt-oss-120b (OpenAI, 2025)² (temperature = 0; prompt hash in the replication package) with a binary *PLAIN USE* vs *QUALIFIED USE* prompt following D7’s rule: explicit textual signals only, reported speech excluded. Of the 527 drawn, 430 remain after stratified balancing, de-

duplication, and removal of occurrences whose term did not survive lexicon curation. On these, D7 and the LLM agree at Cohen’s $\kappa = 0.43$ (and, computed separately on the left- and right-term subsets, $\kappa_{\text{left}} = 0.54$, $\kappa_{\text{right}} = 0.31$), with 118/430 disagreements (27.4%).

The 312 agreement cases and the 118 adjudicated disagreements (each re-examined against the six signals) form the 430-occurrence gold standard. Against this gold, D7 achieves $\kappa = 0.938$ (97.0% agreement), outperforming the LLM judge ($\kappa = 0.49$): D7 deterministically implements the six signals, whereas the LLM drifts into contextual reasoning even when instructed not to. On the 118 adjudicated rows alone, where the gold is not D7’s own output, $\kappa = 0.78$ (89.0%). This validates the operational definition (the same rule applied unchanged to the 2024–2025 corpus), not the broader endorsement construct. As an independent stress-test, a logistic-regression probe on CamemBERT-base embeddings (768-d mean-pool, 5-fold CV) reaches macro F1 = 0.775 against D7’s 0.969 on the same 430 occurrences, and an ensemble of the two underperforms D7 alone (F1 = 0.856): the 110M-parameter contextual model adds no measurable improvement on this operational task. Confusion matrices and UMAP visualisations are in Appendix H.3. Adjudication was by a single author against a written codebook (released in the replication package). The headline $\kappa = 0.938$ therefore measures fidelity to that codebook as applied by one coder, not agreement with an independently constituted gold standard; nor is the LLM comparison independent of it, since the adjudicator resolved the very disagreements the two produced. A second annotator (an unpaid volunteer, fluent in French, with no involvement in the D7 lexicon, rules or pipeline), working from the same codebook, independently relabelled a blind, outlet-masked sample of 100 occurrences stratified over adjudicated and auto-gold cases, agreeing with the author’s gold at $\kappa = 0.94$ (boundary 0.94, control 0.93). On the 13 occurrences where the author’s adjudication overruled the D7 rule, the second annotator reached the author’s label in 12.

For each outlet across all sampled articles in a scoring period:

$$D7_{\text{raw}} = \frac{n_{\text{right}} - n_{\text{left}}}{n_{\text{right}} + n_{\text{left}}}$$

where n_{right} is the count of articles for which plain-

²117B MoE model (Apache 2.0), served via the Regolo API.

use right-term occurrences outnumber qualified-use ones, and n_{left} the analogous count for left terms. A minimum of 10 signal articles is required; otherwise D7 is null for that outlet. Ten is the smallest count at which no single article dominates the majority-vote ratio (shifting $D7_{\text{raw}}$ by at most 0.2 at $n = 10$); it safeguards against sparse outlet samples. We score an all-category corpus pooled over 2024–2025 (568,906 sampled articles, capped at 1,000 per outlet-month); over this window every panel outlet clears the threshold (the thinnest, Mediapart, carries 102 signal articles), giving 30/30 coverage. We score all articles flagged by the curated political lexicon (Appendix H.2) regardless of section tag, since *politique* tagging is inconsistent across outlets. Spearman ρ against the full-model z-scores is 0.93 ($n = 30$); since D7 is itself a fused family, this is an internal-coherence check, not independent validation. Period stability is analyzed in Section 6.3, with per-outlet year-over-year shifts in the replication package.

The D7-only diagnostic submodel is reported in Appendix H. Under the symmetric specification, CNews (+0.15) and Causeur (+0.38) remain clearly separated from Libération (−0.55) and L’Humanité (−0.83), and Le Monde scores center-left (−0.27), its residual undistanced security vocabulary outweighed by its left-coded usage.

6 Results

Because editorial ideology has no observable ground truth, validation relies on convergent validity, robustness, and known-groups discrimination. The question is whether heterogeneous families induce a stable latent ordering that cleanly separates widely recognized ideological groups and persists under perturbation. We do not claim to have recovered the true position of any outlet; we claim a scale that is internally consistent, robust to modeling choices, and convergent with independent external evidence.

Cross-family convergence. Pairwise correlations among the six best-covered families (D0, D3, D4, D6a, D6b, D7; $n \geq 9$ per pair) range from $r = 0.71$ to $r = 0.94$ (mean $r = 0.84$, 15 pairs), consistent with a single latent dimension (see Section 4, Appendix F.3). A formal known-groups test confirms the separation: outlets the external MBFC taxonomy places on the left vs. right do not overlap on the no-D4 holdout scale, which ex-

cludes the grouping signal (AUC = 1.00, Cohen’s $d = 3.05$, exact permutation $p = 0.0006$; Appendix F.2).

6.1 Outlet Ideology Scores

The distribution spans −2.41 (L’Humanité) to +1.87 (Valeurs actuelles): 11 outlets score at or below center-left, 11 score at or above center-right, and 8 are centered. Fifteen outlets in the center region have 95% CIs spanning zero and should be interpreted with caution; the five with $n \leq 4$ direct evidence families (Blast, Le Huff-Post, Slate.fr, Challenges, Atlantico) are the most heavily shrunk toward the prior (for Slate.fr and Challenges, whose prior is zero, toward the panel center; see Appendix D for contribution-share diagnostics).

Figure 1 displays the full ordering with 95% confidence intervals; complete numerical scores are in Appendix Table C.1.

The center-left cluster (20 Minutes through Le Monde, −0.31 to −0.71) is denser than the right tail, a shape that converges with independent surveys of the French press (though partly a function of panel composition; see Limitations). CNews (+0.96) and Le Figaro (+1.02) are effectively tied within the right cluster, with heavily overlapping CIs, consistent with CNews’s rightward shift following Bolloré’s 2015 takeover of its parent network and the subsequent 2017 iTélé–CNews relaunch (Cagé et al., 2024). At the far-right pole, Valeurs actuelles (+1.87) and Fdesouche (+1.84) are nearly indistinguishable: both carry uniformly strong rightward evidence, and VA edges ahead only on its additional ownership family ($D0 = +2.13\sigma$; Fdesouche’s anonymous ownership yields null $D0$). Causeur (+1.53) enters the far-right band only at the margin (barely above the $+1.5\sigma$ cut-point), polemical but less uniformly extreme than the two identitarian outlets.

6.2 Secondary External Corroboration: MBFC Holdout

We hold out the Media Bias Fact Check (D4) evidence family and repurpose it as a secondary external corroboration check, not as the criterion the model is intended to reconstruct. This no-D4 holdout model achieves Spearman $\rho = 0.918$ (95% CI [0.77, 0.97]; $p < 0.001$) against MBFC labels ($n = 15$). Partial correlation among D4, D2, and D6b is expected under convergent validity (Campbell and Fiske, 1959); to rule out circularity, we also

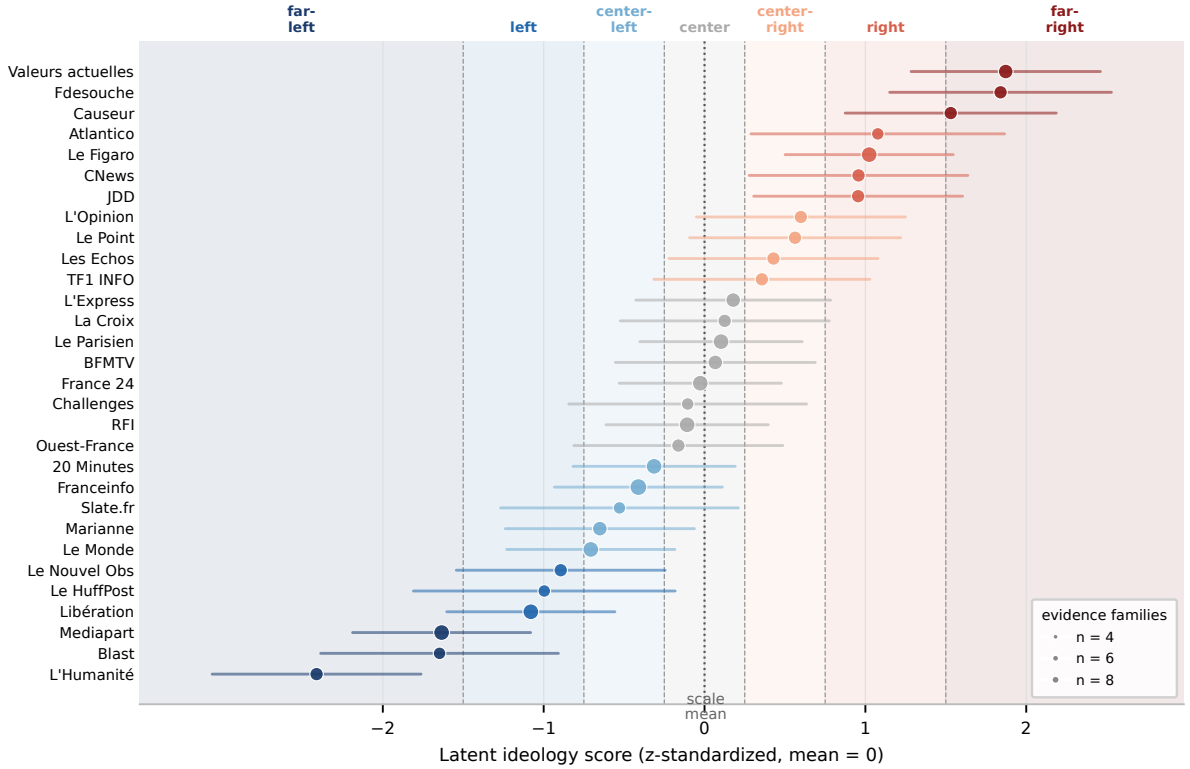


Figure 1: FrIdéo latent ideology scores for 30 French news outlets. Points are full-model continuous scores (z-standardized); whiskers are 95% CIs. Dot size encodes direct evidence families (range: 4–8). Band shading and dashed lines mark σ -band cut-points for the seven interpretive labels (orientation only; all substantive analyses use continuous scores).

compare MBFC against the five families with no methodological overlap (D0, D1, D3, D6a, D7). This structurally independent model agrees at $\rho = 0.931$, essentially matching the full-model value, so the MBFC-overlapping families do not drive agreement. Removing D4 shifts continuous scores at $\rho = 0.9969$ ($n = 30$); maximum shift is 0.224σ (Le HuffPost). Two outlets near σ -band cut-points cross a label boundary (Le Monde, 20 Minutes); continuous scores are preferred downstream.

Ordinal agreement (Table F.1): exact match 6/15, adjacent 8/15, distant 1/15. The sole distant case (Mediapart) reflects a *content–structure gap*: its network position and reader-funded identity align far-left (no-D4 score -1.71) while MBFC’s content-based rater registers its sober investigative style as less partisan (Left-Center).

6.3 Sensitivity Analysis

We drop each of the nine direct families in turn, rerun the full pipeline, and compute Spearman ρ and maximum absolute score shift against the full-model scores (Table B.1). The lowest leave-one-out rank correlation is $\rho = 0.992$ (dropping D6a);

the largest single absolute score shift is 0.27σ (dropping D3, the Cointet network family, at Le Figaro). Dropping D7, the only within-period text family and thus structurally orthogonal to the historical evidence pool, shifts scores by at most 0.25σ (at JDD) yet leaves rank order essentially intact. Across all nine families, LOO $\rho \geq 0.992$; no single family is load-bearing.

Bootstrap stability across 10,000 joint perturbations of all σ_k and $\lambda \in [1.0, 3.0]$ is $\rho = 0.996$ (mean), min 0.988; under uniform $\sigma_k = 0.80$ defaults ranks are unchanged ($\rho > 0.999$). Deterministic refits at $\lambda = 1$ and $\lambda = 3$ move every score $< 0.08\sigma$ ($\rho \geq 0.996$); the sole label change is L’Express, which crosses the center/center-right boundary at $\lambda = 1$ but not $\lambda = 3$.

The distancing layers consistently sharpen the left–right separation (Table B.1, right panel). The guillemet filter is the single largest step and acts asymmetrically: it moves left outlets further left (L’Humanité: $-0.678 \rightarrow -0.822$; Libération: $-0.394 \rightarrow -0.545$), because they cite right-coded vocabulary critically (in reported speech), which bag-of-words (BoW) scoring miscounts as right-

ward; the prefix and symmetric layers add incremental shifts. Right-leaning outlets move only slightly (CNews $+0.225 \rightarrow +0.145$; Causeur $+0.342 \rightarrow +0.423$): their right-coded vocabulary appears predominantly as plain use, so little is filtered. On the all-category corpus the left outlets are already net-negative under BoW: their broad output carries enough plain left-coded usage to set the sign, so the distancing rule sharpens rather than reverses them; its corrective value is largest precisely where right-coded vocabulary is cited critically.

No single term block is determinative: dropping any one of the ten D7 term blocks (the 35 lexical patterns are grouped into five left and five right semantic blocks for this ablation) leaves the fused rank order largely intact ($\rho \geq 0.89$ vs. the full model). The 10-article threshold still matters for narrow sub-periods (e.g. the JDD experiment), where a single half-year can fall below it.

D7 is stable year-over-year when the editorial line is stable, yet moves sharply when it shifts (JDD experiment below). Scoring the all-category corpus separately by calendar year yields $\rho = 0.95$ between the 2024 and 2025 D7 scores ($n = 30$; mean $|\Delta| = 0.09$; per-outlet shifts in the replication package), comparable to the static families. The largest year-over-year shifts are modest ($|\Delta| \leq 0.27$, at Ouest-France) and reflect ordinary corpus variation rather than editorial change; no outlet crosses a label boundary. Contrasted with the sharp JDD discontinuity (Section 7), this confirms D7 moves with genuine editorial change, not sampling noise.

7 Discussion

D7 is an imperfect standalone: security-policy paraphrase (e.g., *violences urbaines* ‘urban violence’ without guillemets) registers identically to plain use when no distancing signal fires, so a center-left outlet can accrue a residual right-coded offset from descriptive usage. D7 is a family within the fusion, not a sufficient statistic.

The JDD natural experiment. The Bolloré overhaul of 22 June 2023 (installing Geoffroy Lejeune, formerly editor of *Valeurs actuelles*, as editorial director in place of Jérôme Béglié) provides a natural experiment with clean before/after structure. $D7_{\text{raw}}$ traces a monotone rightward trajectory tightly aligned with the event (Table 1): net *left* pre-overhaul (consistent with JDD’s es-

tablishment identity), then a $+0.46$ swing in the first six months that flips it to clearly right, climbing further by 2024-H1; all three windows use the identical scoring procedure, so the shift reflects editorial content, not a scoring change. Post-overhaul, JDD treats *immigration massive* (‘mass immigration’) and *ensauvagement* (‘descent into savagery’) as descriptive rather than contested vocabulary, which the plain/qualified rule registers directly. The discontinuity is the paper’s sharpest demonstration of D7’s distinctive property: none of the eight remaining families can register the transition (ownership, citation networks, historical ratings are static within period), whereas D7 detects it temporally.

Table 1: JDD $D7_{\text{raw}}$ trajectory across the 22 June 2023 Bolloré overhaul: the plain-use majority shifts monotonically rightward.

Period	Relation to overhaul	$D7_{\text{raw}}$
Pre-overhaul	Jan 2022 – Jun 2023	-0.160
2023-H2	immediately after	$+0.298$
2024-H1	~ 1 year after	$+0.573$

8 Conclusion

We presented FrIdéo, a continuous, ecosystem-wide ideology scale for 30 French national news outlets, fusing nine heterogeneous evidence families by per-family z-standardization and empirical-Bayes shrinkage toward a founding-tradition prior. Its text family, D7, introduces symmetric plain/qualified scoring, separating plain use of ideologically loaded vocabulary from distanced use on both sides and so reading editorial position from current articles. This makes the scale time-sensitive: D7 registers JDD’s post-2023 realignment, which the eight static families cannot. The ordering survives dropping any single family (LOO $\rho \geq 0.992$) and joint perturbation (bootstrap $\rho = 0.996$), and agrees with held-out MBFC labels. We release the scores, confidence intervals, and replication code.

Limitations

Our source-consistency approach treats disagreement with the remaining evidence pool as a proxy for noise. That assumption is plausible for this applied fusion setting but not guaranteed: a family that captures genuine recent editorial change will disagree with historical consensus sources precisely because it is more informative about current

position. Users who need maximum temporal sensitivity for a specific recently-shifted outlet should weight D7 more heavily or re-run the model with conservative σ_k assignments (tooling provided in the replication package).

The evidence families span 2018 (Pew Research Center, 2018) to 2026 (Media Bias Fact Check, 2026), and French media has undergone significant ownership concentration and editorial drift in that period. CNews’s post-2017 Bolloré transformation and Marianne’s ongoing nationalist pivot are captured imperfectly by the historical families. For outlets undergoing rapid shift within the scoring window (as the JDD analysis demonstrates), D7 should be weighted more heavily or re-run on targeted sub-periods. Missingness handling (Section 4.2) is most defensible for D0 (anonymous or contested ownership) and least so for D6b (Wikipedia profiling), where null could reflect insufficient notability, potentially correlated with outlet age or audience size.

FrIdéo assumes a unidimensional left-right ideology space, consistent with the outlet-scoring literature (Gentzkow and Shapiro, 2010; Budak et al., 2016); PCA on the four highest-coverage families is consistent with a single dominant dimension, including a La Croix/Les Echos test that finds little evidence of a separable economic-vs-cultural second axis (Appendix F.3). Extending FrIdéo to a multidimensional framework (the economic/cultural collapse flagged in Section 4) is a natural direction for future work; candidate subdimension loadings include D0 (ownership → economic axis) and D6b (third-party profiling → cultural axis). The D7 lexical inventory was manually curated using explicit inclusion/exclusion criteria (Appendix H.2) and remains a source of researcher degrees of freedom that can affect scores for sparsely attested outlets (Blast, Slate.fr). Event-driven but ideologically neutral vocabulary is a particular hazard: *émeutiers* (“rioters”), for example, is excluded for neutral saturation, its removal confirmed on the scoring corpus and held-out MBFC (it fires on factual reporting such as the May 2024 Nouvelle-Calédonie unrest rather than on editorial stance). The full-model shrinkage further absorbs residual D7 noise via evidence-pool consensus. Finally, the left-denser shape of the distribution is partly a function of panel composition: a different outlet set would re-define the reference distribution, so we read the distributional shape as convergent with, rather than con-

firmary of, independent surveys. An externally anchored calibration, for example projecting the media scores onto the Chapel Hill Expert Survey (CHES) party-placement scale, would yield an addition-invariant, cross-panel-comparable scale and would quantify the systematic rightward offset of the media centre relative to the party spectrum. We regard this as the primary planned extension but leave it to future work, since it requires an external bridge between party and media placements that the present evidence families do not supply.

Reproducibility Statement

The core scoring pipeline (z-standardization, empirical-Bayes shrinkage, bootstrap, LOO) depends only on the Python standard library; PCA (unidimensionality validation) and the optional GMM label-assignment path use scikit-learn, while the D7 lexical scoring, shrinkage estimation, and all sensitivity analyses require no external ML framework. The bootstrap random seed is fixed at 42. Full replication code, outlet evidence data, scored results, and a pinned dependency lockfile are available at <https://github.com/lefrenchnewslab/frideo>.

References

- Eytan Bakshy, Solomon Messing, and Lada A. Adamic. 2015. [Exposure to ideologically diverse news and opinion on facebook](#). *Science*, 348(6239):1130–1132.
- Rodney Benson and Daniel C. Hallin. 2007. [How states, markets and globalization shape the news: The french and US national press, 1965–97](#). *European Journal of Communication*, 22(1):27–48.
- Ceren Budak, Sharad Goel, and Justin M. Rao. 2016. [Fair and balanced? quantifying media bias through crowdsourced content analysis](#). *Public Opinion Quarterly*, 80(S1):250–271.
- Julia Cagé. 2017. [Qui possède les médias ?](#) Technical Report 33, LIEPP Policy Brief.
- Julia Cagé, Moritz Hengel, Nicolas Hervé, and Camille Urvoy. 2022. [Hosting media bias: Evidence from the universe of french broadcasts, 2002–2020](#). Technical Report hal-03878119, Sciences Po. Preprint, version of 16 February 2022.
- Julia Cagé, Moritz Hengel, Nicolas Hervé, and Camille Urvoy. 2024. [Hosting media bias: Evidence from the universe of french broadcasts, 2002–2020](#). Technical Report 537, CRC TR 224 Discussion Paper Series. First version: May 2021; this version: March 2024.

- Donald T. Campbell and Donald W. Fiske. 1959. [Convergent and discriminant validation by the multitrait-multimethod matrix](#). *Psychological Bulletin*, 56(2):81–105.
- Jean-Philippe Cointet, Dominique Cardon, Andreï Mogoutov, Benjamin Ooghe-Tabanou, Guillaume Plique, and Pedro Ramaciotti Morales. 2021. [Uncovering the structure of the french media ecosystem](#). *Preprint*, arXiv:2107.12073. IC2S2 2021.
- Sebastián Concha Macías and Brian Keith Norambuena. 2024. [Evaluating the ability of computationally extracted narrative maps to encode media framing](#). In *Proceedings of the 7th International Workshop on Narrative Extraction from Texts (Text2Story@ECIR 2024)*.
- Rohan Das, Aditya Chandra, I-Ta Lee, and Maria Leonor Pacheco. 2024. [Media framing through the lens of event-centric narratives](#). In *Proceedings of the 6th Workshop on Narrative Understanding*, pages 85–98, Miami, Florida, USA.
- Bradley Efron and Carl Morris. 1977. [Stein’s paradox in statistics](#). *Scientific American*, 236(5):119–127.
- Fondation Descartes. 2020. [Photographie du nouveau paysage politico-médiatique français](#). Report.
- Matthew Gentzkow and Jesse M. Shapiro. 2010. [What drives media slant? evidence from U.S. daily newspapers](#). *Econometrica*, 78(1):35–71.
- Tim Groseclose and Jeffrey Milyo. 2005. [A measure of media bias](#). *The Quarterly Journal of Economics*, 120(4):1191–1237.
- Felix Hamborg, Karsten Donnay, and Bela Gipp. 2019. [Automated identification of media bias in news articles: An interdisciplinary literature review](#). *International Journal on Digital Libraries*, 20(4):391–415.
- Simon Jackman. 2001. [Multidimensional analysis of roll call data via bayesian simulation: Identification, estimation, inference, and model checking](#). *Political Analysis*, 9(3):227–241.
- Willard James and Charles Stein. 1961. [Estimation with quadratic loss](#). In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 361–379, Berkeley, CA. University of California Press.
- Eunji Kim, Yphtach Lelkes, and Joshua McCrain. 2022. [Measuring dynamic media bias](#). *Proceedings of the National Academy of Sciences*, 119(32):e2202197119.
- Sora Lim, Adam Jatowt, Michael Färber, and Masatoshi Yoshikawa. 2020. [Annotating and analyzing biased sentences in news articles using crowdsourcing](#). In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 1478–1484, Marseille, France. European Language Resources Association.
- Gregory J. Martin and Ali Yurukoglu. 2017. [Bias in cable news: Persuasion and polarization](#). *American Economic Review*, 107(9):2565–2599.
- Media Bias Fact Check. 2026. [Media bias fact check source ratings](#). Ratings pages and API documentation accessed 2026-05-31.
- OECD and European Commission Joint Research Centre. 2008. [Handbook on Constructing Composite Indicators: Methodology and User Guide](#). OECD Publishing, Paris.
- OpenAI. 2025. [gpt-oss-120b & gpt-oss-20b model card](#). *Preprint*, arXiv:2508.10925.
- Pew Research Center. 2018. [News media and political attitudes in france](#). Fact sheet.
- Franck Rebillard and Jedediah Sklower. 2022. [Le pluralisme des médias à l’ère numérique](#). Technical Report hal-03723303, Centre for Media Pluralism and Media Freedom; Robert Schuman Centre for Advanced Studies; European University Institute; Iméccen. Research report.
- Ronja Rönnback, Chris Emmery, and Henry Brighton. 2025. [Automatic large-scale political bias detection of news outlets](#). *PLOS ONE*, 20(5):e0321418.
- Boris Shor and Nolan McCarty. 2011. [The ideological mapping of american legislatures](#). *American Political Science Review*, 105(3):530–551.
- Jin Zhao, Xinrui Hu, and Nianwen Xue. 2025. [Seeing the same story differently: Framing-Divergent event coreference for computational framing analysis](#). In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 28356–28371, Suzhou, China.

A Evidence Families

Table A.1 specifies the ten evidence inputs of Section 4—source, construct measured, source-consistency parameter σ_k , and outlet coverage. Coverage ranges by design from 3/30 (D1, TV only) to 30/30 (D3, D7, S2); it is this variation the shrinkage step absorbs.

B Sensitivity and Distancing-Ablation Tables

Table B.1 supports Section 6.3. Its left panel reports the leave-one-family-out analysis (drop each direct family, refit, compare to the full model); its right panel isolates each distancing layer in D7 across five high-signal outlets, from raw bag-of-words to the full symmetric check.

Table A.1: Evidence families used in the ideology scoring framework.

ID	Source	What it measures	Source consistency parameter σ_k	Outlet coverage
D0	Acrimed/ <i>Le Monde diplo.</i> ownership map (via data.gouv.fr); press/Wikipedia for absent or recently-sold outlets	Owner political ideology via structural map	0.35	28/30 (null: concealed or ideologically-unmarked owner)
D1	Cagé et al. (2022)	Guest radical-right deviation relative to France 2 benchmark (pp)	0.57	3/30 (TV channels)
D2	Pew Research Center (2018)	Audience-perceived ideology (0–6 scale; 3 = center), directly-published placement means	0.57	8/30
D3	Cointet et al. (2021)	Our continuous position via dual PageRank over Cointet’s web-link network (421 nodes, 20,992 edges), seeded on its community blocks; right/left log-ratio	0.63	30/30 (28 PPR + 2 categorical fallback)
D4	Media Bias Fact Check (2026)	Categorical bias label (far-left to far-right)	0.50	15/30
D6a	Editorial charters / statutes	Self-declared political orientation	0.48	18/30
D6b	Wikipedia / press typology	Third-party journalistic profiling	0.30	27/30
D7	This work (signal-based lexical)	Content-based lexical ideology: plain-use occurrences outnumber qualified-use (signal-present) occurrences	0.42	30/30
D8	Ad Fontes Media	Categorical bias label from independent US rater (same ordinal scale as D4; normalised via D4 reference distribution)	0.24	9/30
S2	Historical record	Founding editorial tradition (5-point scale)	prior only	30/30

Table B.1: **Left:** LOO sensitivity — rank correlation and maximum absolute score shift when each direct family is dropped in turn. **Right:** D7 distancing ablation for five high-signal outlets (2024–2025); BoW = raw bag-of-words, Guill. = guillemet filter, G+P = adds distancing prefix, Full = full symmetric (adds side-specific critique markers); Full differs from Table H.1 by <0.05 (per-year vs. pooled).

Family	n	LOO ρ	Max $ \Delta $	Outlet	Outlet	BoW	Guill.	G+P	Full
D0 (ownership)	28	0.9973	0.196	Libération	L’Humanité	−0.678	−0.822	−0.830	−0.836
D1 (Cagé guests)	3	0.9969	0.145	Le Figaro	Libération	−0.394	−0.545	−0.547	−0.550
D2 (Pew survey)	8	0.9973	0.182	Le Monde	Le Monde	−0.196	−0.297	−0.290	−0.289
D3 (Cointet PPR)	30	0.9956	0.268	Le Figaro	CNews	+0.225	+0.149	+0.147	+0.145
D4 (MBFC)	15	0.9969	0.224	Le HuffPost	Causeur	+0.342	+0.410	+0.416	+0.423
D6a (editorial charters)	18	0.9920	0.207	JDD					
D6b (Wikipedia profiling)	27	0.9969	0.165	Le Figaro					
D7 (signal-based lexical)	30	0.9933	0.250	JDD					
D8 (Ad Fontes Media)	9	1.0000	0.059	TF1 INFO					

C Full-Model Outlet Ideology Scores

Table C.1 is the released resource: final score, analytic standard error, 95% CI, seven-band label, and direct-family count for all 30 outlets. Scores are zero-mean and unit-variance across the panel by construction (Section 4.3), hence interpretable only relative to this outlet set.

D Normalized Additive Contribution Shares

Figure D.1 decomposes each final score into signed additive contributions from five evidence groups, defined by methodological provenance: Ownership (D0), Academic & network (D1, D2, D3), Ratings (D4, D8), Editorial & content (D6a, D6b, D7), and the Tradition prior (S2). For outlet i , group g contributes $C_{ig} = \sum_{k \in g} C_{ik}$, with

$C_{ik} = (\alpha_i/n_i) \cdot z_{ik}/\sigma_B$ for direct families and $C_{i,\text{prior}} = (1 - \alpha_i) \cdot z_i^{\text{struct}}/\sigma_B$ for the prior, on the final standardized scale (SD units). These are *normalized additive decomposition shares*, not causal attributions or Shapley values.

Three patterns emerge. Evidence-rich outlets (Le Monde $n = 7$, Franceinfo $n = 8$, Valeurs actuelles $n = 6$) show balanced multi-group support, the prior contributing under 25% of absolute signal mass. Sparse outlets (Blast, Slate.fr, both $n = 4$) are dominated by at most two groups, with roughly a third of shrinkage weight $(1 - \alpha)$ on the prior; where $S2 = 0$ (Slate.fr) it contributes none and shrinkage pulls toward the panel center. For near-center outlets (BFMTV, Ouest-France) positive and negative contributions offset: a small score reflects genuine evidence tension, not absent signal. Full family-level decomposition: replica-

Table C.1: Full-model ideology scores, all 30 outlets, two panels. Label is for interpretability only; all analyses use the continuous Score. The analytic σ_i is recoverable to within 0.01 as one-quarter of the CI width divided by 0.98. [†]CI spans zero (15 outlets, all near-center). Five outlets have $n \leq 4$ direct families (Blast, Le HuffPost, Slate.fr, Challenges, Atlantico): shrinkage places roughly a third of the weight on the tradition prior ($\alpha = 0.67$ at $n = 4$), so these are heavily shrunk and their labels indicative only. For Slate.fr and Challenges the prior is zero ($S2 = 0$), contributing no signed mass; shrinkage instead pulls them toward the panel center. See Appendix D.

Outlet	Score	95% CI	Label	n	Outlet	Score	95% CI	Label	n
L’Humanité	−2.41	[−3.06, −1.76]	far-left	5	BFMTV [†]	+0.07	[−0.55, +0.69]	center	6
Blast	−1.65	[−2.39, −0.91]	far-left	4	Le Parisien [†]	+0.10	[−0.40, +0.61]	center	7
Mediapart	−1.63	[−2.19, −1.08]	far-left	7	La Croix [†]	+0.13	[−0.52, +0.77]	center	5
Libération	−1.08	[−1.60, −0.56]	left	7	L’Express [†]	+0.18	[−0.43, +0.78]	center	6
Le HuffPost	−1.00	[−1.81, −0.18]	left	4	TF1 INFO [†]	+0.36	[−0.31, +1.03]	center-right	5
Le Nouvel Obs	−0.89	[−1.54, −0.24]	left	5	Les Echos [†]	+0.43	[−0.22, +1.08]	center-right	5
Le Monde	−0.71	[−1.23, −0.18]	center-left	7	Le Point [†]	+0.56	[−0.09, +1.22]	center-right	5
Marianne	−0.65	[−1.24, −0.06]	center-left	6	L’Opinion [†]	+0.60	[−0.05, +1.25]	center-right	5
Slate.fr [†]	−0.53	[−1.27, +0.21]	center-left	4	JDD	+0.95	[+0.31, +1.60]	right	5
Franceinfo [†]	−0.41	[−0.93, +0.11]	center-left	8	CNews	+0.96	[+0.28, +1.64]	right	5
20 Minutes [†]	−0.31	[−0.82, +0.19]	center-left	7	Le Figaro	+1.02	[+0.50, +1.55]	right	7
Ouest-France [†]	−0.16	[−0.81, +0.49]	center	5	Atlantico	+1.08	[+0.29, +1.87]	right	4
RFI [†]	−0.11	[−0.61, +0.40]	center	7	Causeur	+1.53	[+0.87, +2.19]	far-right	5
Challenges [†]	−0.10	[−0.84, +0.63]	center	4	Fdesouche	+1.84	[+1.15, +2.53]	far-right	5
France 24 [†]	−0.03	[−0.53, +0.48]	center	7	Valeurs actuelles	+1.87	[+1.28, +2.46]	far-right	6

tion package.

E Post-2019 Shifters and the No-Prior Refit

Which family corrects a stale D3? Decomposing the final score into signed family contributions (Appendix D) gives, in SD units and % of absolute mass: **JDD** — D7 +0.355 (32%), D0 +0.286 (26%), D6b +0.260 (24%); **CNews** — D0 +0.286 (29%), D6b +0.260 (27%), D1 +0.241 (25%), D7 +0.152 (16%). D7 leads at the ownership change it was designed to catch, but is not the general corrective.

This is also why LOO understates D7: dropping it moves JDD only 0.25σ (Table B.1) despite supplying 32% of its mass, because D0, D6b and the prior pull the same way. LOO measures marginal displacement under corroboration, not share of evidence—D6a’s largest shift is likewise at JDD (0.207).

No-prior refit. Refitting with $\alpha_i = 1$ throughout (prior removed, then re-standardized) tests whether the face-validity agreement is circular. All three prior-dependent sparse outlets keep their labels: Blast $-1.65 \rightarrow -1.65$, Le HuffPost $-1.00 \rightarrow -1.04$, Atlantico $+1.08 \rightarrow +1.16$. Only two labels change panel-wide, neither sparse: L’Express (center $\rightarrow$ center-right) and Causeur (far-right $\rightarrow$ right), both already flagged as cut-point sensitive.

F Validation

This appendix collects the three validity checks of Section 6: convergent agreement with a held-out rating source (F.1), separation of externally-defined known groups (F.2), and a test of the unidimensionality assumption (F.3). All three use hold-out or out-of-model quantities, so the criterion is never also an input.

F.1 MBFC Holdout: Outlet-Level Comparison

Table F.1 compares labels from the *no-D4 holdout model*—refitted with the MBFC family removed—against MBFC’s own labels for the 15 outlets it covers. Since D4 is excluded from the model being evaluated, agreement is convergent corroboration, not reconstruction of a fitted target.

F.2 Known-Groups Separation

Do outlets an *external* taxonomy assigns left vs. right separate on the FrIdéo scale? Groups follow MBFC bands (left = *Far Left/Left-Center*; right = *Right-Center/Right/Far Right*; the two *Center* outlets are reported but excluded). Scores are the *no-D4 holdout* scores, so the grouping variable is not an input to what it separates. Separation is complete: every right-group outlet outscores every left-group outlet. The left group ($n = 7$) averages -0.80 (range -1.71 to -0.20), the right ($n = 6$) $+1.14$ ($+0.00$ to $+1.89$), the two Center

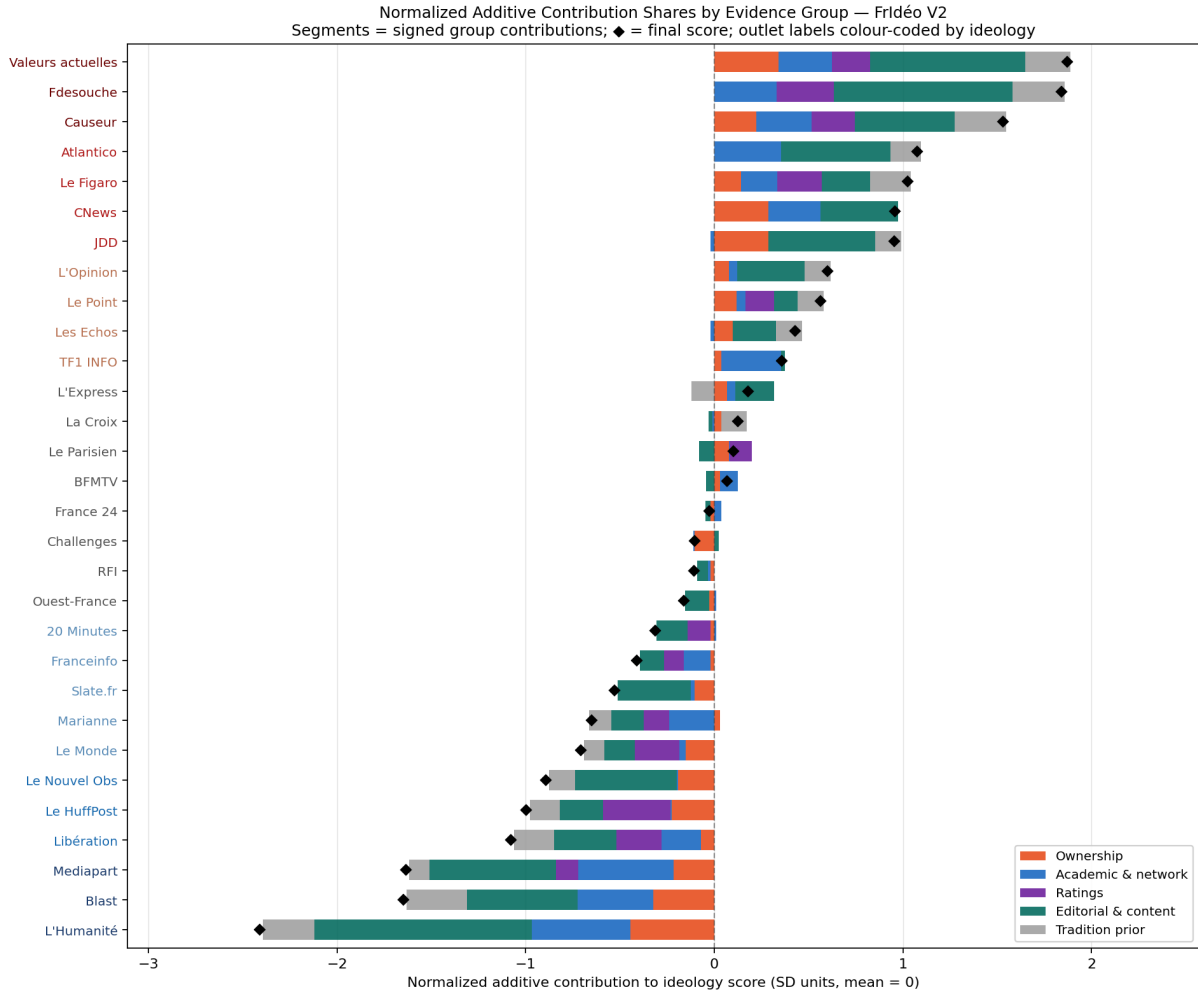


Figure D.1: Normalized additive contribution shares by evidence group, all 30 outlets (spectrum left to right, bottom to top). Each segment sums signed final-scale contributions (C_{ig} , SD units) within a group; the diamond (♦) marks the final score. Groups: **orange** Ownership (D0); **blue** Academic & network (D1–D3); **purple** Ratings (D4, D8); **teal** Editorial & content (D6a, D6b, D7); **grey** Tradition prior (S2).

outlets -0.05 . Mann–Whitney $U = 42/42$ (AUC = 1.00), Cohen’s $d = 3.05$, exact one-sided permutation $p = 0.0006$ (1/1716).

F.3 Unidimensionality Validation

We test by PCA whether the space is unidimensional and whether a second dimension is economic-versus-cultural. Inputs are the four highest-coverage families (D0, D3, D6b, D7; each ≥ 27 outlets) over the 25 complete-case outlets, z-standardized as in the main pipeline, eigen-decomposed without rotation.

Results. PC1 explains 88.5% of shared variance, PC2 7.4% (scree ratio 11.9 $\times$). All four families load strongly on PC1 (D0 -1.02 , D6b -0.97 , D7 -0.95 , D3 -0.85), consistent with one dominant latent dimension. PC2 is driven mainly by D3 (-0.49), the rest loading weakly ($|\text{loading}| \leq 0.18$)

— little evidence the residual axis is economic-versus-cultural. Scree plot and biplot: replication package.

Economic vs. cultural test. If PC2 separated economic conservatism (D0) from cultural conservatism (D6b), La Croix and Les Echos — which reach right-of-center by different routes — should sit at opposite PC2 poles. They differ by only 0.35 SD (La Croix -0.10 , Les Echos $+0.25$), and D3 loads opposite to D0 on PC2 (-0.49 vs. $+0.14$): network position and ownership capture partially distinct signals, not a clean economic/cultural split.

G Structural Prior (S2) Coding

S2 is the prior, not a direct family. It encodes *founding* editorial tradition only — deliberately distinct from current ownership (D0) and cur-

Table F.1: No-D4 convergent validity vs. MBFC. Model = *no-D4 holdout* labels, not full-model labels (e.g. Le Monde is center-left in the full model, left here).

Outlet	Score	Model	MBFC	Agree
Mediapart	−1.71	far-left	Left-Center	× (2 cat.)
Libération	−1.20	left	Left-Center	~ (adja-cent)
Le Monde	−0.79	left	Left-Center	~ (adja-cent)
Le Huff-Post	−0.77	left	Far Left	~ (adja-cent)
Marianne	−0.60	center-left	Left-Center	✓
Franceinfo	−0.35	center-left	Left-Center	✓
20 Minutes	−0.20	center	Left-Center	~ (adja-cent)
RFI	−0.10	center	Center/Least Biased	✓
France 24	−0.01	center	Center/Least Biased	✓
Le Parisien	+0.00	center	Right-Center	~ (adja-cent)
Le Point	+0.47	center-right	Right-Center	✓
Le Figaro	+1.18	right	Right-Center	~ (adja-cent)
Causeur	+1.50	far-right	Right	~ (adja-cent)
Fdesouche	+1.77	far-right	Far Right	✓
Valeurs actuelles	+1.89	far-right	Right	~ (adja-cent)

rent self-declaration or profiling (D6a/D6b) — so it reflects historical baseline, not present-day line. Authors assigned values on a symmetric five-point grid $\{-0.5, -0.25, 0, +0.25, +0.5\}$ from each outlet’s documented founding profile (year, founder or institution, initial alignment) using published press histories; per-outlet sources are in the replication package. The ± 0.5 cap is by design: the prior is weak relative to direct evidence and is down-weighted by shrinkage $(1 - \alpha_i)$ as families accrue, so it anchors only data-sparse outlets and is overridden wherever direct families exist (JDD’s +0.25 founding tradition is dominated by its post-2023 D6b signal). Coverage 30/30.

H D7 Lexical Scoring: Full Results and Term Selection

This appendix reports the D7 family in full: per-outlet scores (H.1), lexicon construction and rejection log (H.2), and the embedding-probe stress test (H.3).

H.1 Outlet-Level D7 Scores

Table H.1 gives $D7_{\text{raw}}$ for all 30 outlets on the all-category 2024–2025 corpus, with the underlying

left- and right-signal article shares and per-outlet scored-article counts. Every outlet clears the 10-signal-article threshold on this window (30/30 coverage).

H.2 Term Selection Log

From roughly 80 candidates, terms were culled manually against three operationalizable criteria; no automated correlation search was run and no terms were added post-hoc. *Neutral saturation*: fires at high, roughly equal rates spectrum-wide, journalistically unmarked. *Polysemy*: domain-specific use (religious, legal) fires reliably in ideologically opposed outlets. *Wire contamination*: enters opposing outlets via AFP/official-language import, not editorial endorsement. Retained terms were validated by endorsed firing rates across five left- and six right-cluster outlets on the 2024–2025 corpus (median discrimination $4.6\times$; full per-term ratios in the replication package). Raw occurrence rates are not used: they conflate plain and distanced use.

Rejected terms. Grouped by rejection criterion, with the outlet at which the term fired at high rate despite an opposing editorial stance:

- *Event saturation*. *émeutiers* (‘rioters’)—coverage-heavy outlets during the May 2024 Nouvelle-Calédonie unrest; lexically neutral.
- *Neutral saturation*. *délinquance* (‘delinquency’), *criminalité* (‘criminality’)—Le Monde, La Croix (factual social reporting); *flux migratoires* (‘migratory flows’)—all outlets (demographic register); *souveraineté nationale* (‘national sovereignty’), *identité nationale* (‘national identity’)—spectrum-wide (constitutional/EU discourse); *sans-papiers* (‘undocumented migrants’), *évasion fiscale* (‘tax evasion’)—Les Echos, Ouest-France (business/civic reporting).
- *Polysemy*. *islamisation* (‘Islamization’), *séparatisme* (‘separatism’), *communautarisme* (‘communitarianism’)—La Croix (religious-affairs journalism); *assistan[ct]*—all outlets (*assistant social* ‘social worker’ collocation).
- *Wire contamination*. *clandestins* (‘illegal migrants’)—La Croix (AFP wire / official language).

H.3 Signal Validation: Embedding Probe

Figure H.1 compares a logistic-regression probe on CamemBERT-base embeddings against D7

Table G.1: Complete S2 founding-tradition coding for all 30 outlets, in two panels. Value is the ordinal S2 input (scaled by 1.5 in the model). Coverage 30/30.

Outlet	S2	Founding basis	Outlet	S2	Founding basis
L'Humanité	−0.5	1904, Jaurès — PCF organ	Ouest-France	0.0	1944 (SIPA) — independent
Blast	−0.5	2021, D. Robert — left investigative	RFI	0.0	1975 — public international radio
Libération	−0.5	1973, Sartre — radical left	France 24	0.0	2006 — public international news
Mediapart	−0.25	2008, Plenel — investigative, left	Challenges	0.0	1982 — centrist business weekly
Le HuffPost	−0.25	2012, French HuffPost — progressive	CNews	0.0	i-Télé 1999 — centrist; shift via D1
Le Monde	−0.25	1944, Beuve-Méry — centre-left	La Croix	+0.25	1880, Assumptionists — Catholic
Le Nouvel Obs	−0.25	1964, Daniel — non-communist left	Les Echos	+0.25	1908 — liberal business daily
Marianne	−0.25	1997, J.-F. Kahn — left-republican	Le Point	+0.25	1972, Imbert — Gaullist right
L'Express	−0.25	1953, Servan-Schreiber — centre-left	JDD	+0.25	1948, Lazareff — centre-right; pre-Bolloré
Slate.fr	0.0	2009, French offshoot of US Slate	Atlantico	+0.25	2011, Ferjou — liberal right
Franceinfo	0.0	FTV/Radio France — neutral remit	L'Opinion	+0.25	2013, Beytout — pro-business
BFMTV	0.0	2005 — commercial 24h news	Le Figaro	+0.5	1826 — conservative
TF1 INFO	0.0	TF1 digital; privatised 1987	Causeur	+0.5	2007, É. Lévy — conservative-right
20 Minutes	0.0	2002 — free mass-distribution daily	Valeurs actuelles	+0.5	1966 — far-right
Le Parisien	0.0	1944 — popular Paris daily	Fdesouche	+0.5	2006 — far-right aggregator

Table H.1: D7 lexical scores, all 30 outlets (2024–2025 corpus), two panels. $D7_{\text{raw}} = (n_{\text{right}} - n_{\text{left}})/(n_{\text{right}} + n_{\text{left}})$; L% and R% are shares of articles whose plain-use occurrences for that side outnumber qualified-use ones. n_art. counts scored articles; articles with unusable body text are dropped, so the column sums just below the 568,906 sampled total.

Outlet	n_art.	L%	R%	D7_raw	Outlet	n_art.	L%	R%	D7_raw
L'Humanité	24,367	6.3%	0.6%	−0.829	Challenges	16,980	1.5%	1.1%	−0.145
Mediapart	9,234	1.0%	0.1%	−0.784	Le Parisien	25,000	0.6%	0.5%	−0.137
Libération	25,000	1.9%	0.5%	−0.547	TF1 INFO	25,000	1.1%	0.9%	−0.110
Blast	959	12.5%	3.8%	−0.538	Le Point	24,224	1.3%	1.1%	−0.098
Le HuffPost	19,841	1.8%	0.6%	−0.535	RFI	24,358	1.3%	1.1%	−0.078
Slate.fr	5,543	2.6%	0.8%	−0.518	L'Express	18,805	2.3%	2.3%	−0.005
Le Nouvel Obs	24,004	1.8%	0.6%	−0.471	France 24	10,408	2.4%	2.4%	+0.002
20 Minutes	25,000	0.9%	0.4%	−0.357	L'Opinion	14,913	0.8%	0.9%	+0.079
Ouest-France	25,000	0.3%	0.1%	−0.346	Le Figaro	25,000	0.5%	0.6%	+0.144
Marianne	12,196	2.2%	1.1%	−0.320	CNews	25,000	0.6%	0.9%	+0.147
La Croix	25,000	1.6%	0.8%	−0.316	Atlantico	6,411	3.7%	5.6%	+0.195
Franceinfo	25,000	1.6%	0.9%	−0.290	Causeur	4,747	5.0%	11.1%	+0.375
Le Monde	25,000	1.1%	0.6%	−0.274	Valeurs actuelles	19,258	1.3%	4.5%	+0.540
BFMTV	25,000	1.2%	0.7%	−0.246	JDD	22,036	1.5%	5.1%	+0.546
Les Echos	25,000	1.1%	0.8%	−0.169	Fdesouche	10,613	0.8%	2.6%	+0.554

rules on the 430-occurrence gold standard (5-fold stratified CV). D7 reaches macro F1 = 0.969 with precision = 0.988 for PLAIN. The probe reaches 0.775, and an ensemble underperforms D7 alone (0.856). The confusion matrices show why: D7's errors are predominantly conservative (10 false QUALIFIED against 3 false PLAIN), the probe's bidirectional. UMAP corroborates this — no term, retained or rejected, shows a clean outlet-ideology axis (per-term $|\rho| \lesssim 0.4$). Correlations and UMAP plots: replication package.

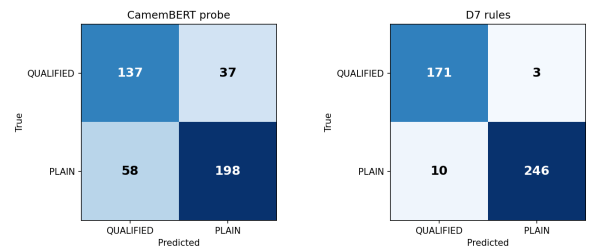


Figure H.1: Confusion matrices for the CamemBERT linear probe (left) and D7 rules (right) on the 430-occurrence PLAIN/QUALIFIED gold standard.