

Secularized Purity: Cross-Platform Evidence for Implicit Use of Binding Moral Foundations in Progressive Italian Social Media Discourse

Franco Cazzaniga
Università degli Studi dell’Insubria
franco.cazzaniga@uninsubria.it

January 11, 2026

Abstract

Moral Foundations Theory (MFT) posits that political conservatives rely on all five moral foundations (Care, Fairness, Loyalty, Authority, Sanctity), whereas progressives primarily endorse the individualizing foundations of Care and Fairness. This study examines whether this asymmetry characterizes actual moral language use in naturalistic political discourse. Analyzing 4,789 Italian-language social media posts from Mastodon ($N = 2,812$) and Reddit ($N = 1,977$), we investigate the implicit deployment of moral foundations in progressive discourse using a custom Italian lexicon and robust non-parametric statistics. Across both platforms, 85–90% of progressive-coded posts contain markers of binding foundations, with large within-group effect sizes (Cramér’s $V = 0.71$ – 0.80). Authority-related language appears significantly more frequently in progressive discourse on both platforms, while Sanctity-related contamination rhetoric shows consistent directional effects with small effect sizes. These findings suggest that the liberal–conservative asymmetry described by MFT reflects differences in explicit moral endorsement rather than underlying moral-cognitive structure. Progressive discourse systematically deploys binding moral foundations through a secularized vocabulary of authority and moral contamination, despite explicit commitments to individualizing foundations alone.

Keywords: Moral Foundations Theory, political psychology, social media analysis, progressive discourse, purity rhetoric, cross-platform validation, non-parametric statistics

1 Introduction

Jonathan Haidt’s Moral Foundations Theory (MFT) has become one of the most influential frameworks for understanding political-moral psychology (Graham et al., 2009; Haidt, 2012). The theory proposes five (later six) innate moral “taste receptors”: Care/Harm, Fairness/Cheating, Loyalty/Betrayal, Authority/Subversion, and Sanctity/Degradation. A central empirical claim of MFT is that political liberals rely predominantly on the “individualizing” foundations (Care and Fairness), while conservatives draw more equally from all five foundations, including the “binding” foundations (Loyalty, Authority, Sanctity) (Graham et al., 2009).

This asymmetry has been replicated across multiple studies using self-report measures (Graham et al., 2013), text analysis of sermons (Graham et al., 2009), and the Moral Foundations Questionnaire (Graham et al., 2011). However, most research has measured *explicit endorsement* of moral values rather than their *implicit usage* in naturalistic discourse.

We propose that progressives may implicitly employ all five moral foundations while explicitly endorsing only Care and Fairness. Specifically, we hypothesize that:

1. Progressive discourse implicitly invokes Loyalty (in-group/out-group distinctions, movement solidarity)

2. Progressive discourse implicitly invokes Authority (appeals to science, institutions, expertise)
3. Progressive discourse implicitly invokes Sanctity (moral contamination labels like “fascist,” “racist,” “toxic”)
4. These patterns are robust across different social media platforms

This study tests these hypotheses using computational text analysis of Italian political discourse on two distinct platforms: Mastodon (a decentralized microblogging platform) and Reddit (a forum-based discussion platform). The cross-platform design provides a crucial test of robustness.

2 Literature Review

2.1 Moral Foundations Theory

MFT emerged from cross-cultural research on moral reasoning (Shweder et al., 1997; Haidt and Joseph, 2004). Haidt and Kesebir (2007) proposed that human morality is built on multiple innate foundations that are then culturally elaborated. The theory distinguishes between:

- **Individualizing foundations:** Care/Harm (compassion, preventing suffering) and Fairness/Cheating (justice, rights, reciprocity)
- **Binding foundations:** Loyalty/Betrayal (group cohesion, patriotism), Authority/Subversion (respect for hierarchy, tradition), and Sanctity/Degradation (purity, contamination avoidance)

The seminal study by Graham et al. (2009) found that liberals consistently prioritized Care and Fairness, while conservatives endorsed all five foundations more equally. This finding has been interpreted as explaining why liberals and conservatives “talk past each other”—they are operating with different moral vocabularies (Haidt, 2012).

2.2 Critiques and Limitations

Several scholars have questioned the liberal-conservative asymmetry. Frimer et al. (2020) found that in certain contexts (e.g., climate change discourse), liberals used Authority and Purity language as much as or more than conservatives. Turner-Zwinkels et al. (2021) showed that conservatives’ moral foundations are more densely interconnected, suggesting qualitative rather than merely quantitative differences.

Importantly, most MFT research relies on self-report measures, which may capture *explicit beliefs* rather than *implicit moral reasoning*. Jost and Amodio (2009) argue that moral foundations may reflect post-hoc rationalizations of ideological positions rather than their causes.

2.3 Computational Approaches to MFT

The development of the Moral Foundations Dictionary (MFD) (Graham et al., 2009) and its successors (Frimer et al., 2019; Hopp et al., 2021) has enabled large-scale text analysis of moral language. Studies have applied these tools to political tweets (Johnson and Goldwasser, 2018), news media (Mokhberian et al., 2020), and online discourse (Hoover et al., 2020).

Recent work has examined MFT in non-English contexts, including Italian social media discourse on immigration (Lai et al., 2021). However, no study has systematically examined whether progressives *implicitly* use binding foundations that they would not *explicitly* endorse, and none have validated findings across multiple platforms.

2.4 The Secularization of Purity

A key theoretical contribution of this paper is the concept of “secularized purity rhetoric.” Traditional Sanctity/Degradation language involves religious or bodily contamination (e.g., “unclean,” “sinful,” “disgusting”). We propose that progressive discourse has developed a parallel vocabulary of *moral contamination* using political-ideological labels.

Terms like “fascist,” “racist,” “toxic,” and “problematic” function as contamination markers—they designate individuals or ideas as morally polluted and requiring exclusion from the moral community. This parallels the conservative use of terms like “degenerate” or “perverted,” but targets different objects.

3 Methods

3.1 Data Collection

We collected data from two distinct platforms to ensure robustness:

3.1.1 Mastodon (N=2,812)

Posts were collected from Italian Mastodon instances (mastodon.uno, sociale.network, mastodon.bida.im) using public API endpoints. Collection included:

- Public timeline sampling
- Political hashtag searches (#politica, #italia, #governo, #sinistra, #destra, #antifascismo, #immigrazione, etc.)

3.1.2 Reddit (N=1,977)

Posts were collected from Italian subreddits (r/italy, r/italia) using Reddit’s public JSON API. Collection included:

- Hot and new post sampling
- Political keyword searches (politica, governo, meloni, salvini, fascismo, antifascismo, immigrazione, etc.)

Posts were filtered for minimum length (30 characters). The final corpus comprised 4,789 documents.

3.2 Lexicon Development

We developed a custom Italian MFT lexicon containing 337 moral markers across five foundations. Unlike existing dictionaries, our lexicon distinguishes markers by *political orientation* (progressive, populist, neutral), allowing us to track foundation usage within ideological camps.

Key innovations include:

- **Progressive Sanctity markers:** “fascista,” “razzista,” “sessista,” “omofobo,” “tossico/a,” “problematico/a”
- **Progressive Authority markers:** “la scienza,” “gli esperti,” “consenso scientifico,” “fact-checking,” “Europa,” “OMS”
- **Progressive Loyalty markers:** “comunità,” “alleato/a,” “movimento,” “dalla parte giusta della storia”

3.3 Classification Procedure

Each post was analyzed for the presence of lexicon markers. Posts were classified as “progressive” or “populist” based on the predominant orientation of detected markers. Posts with no markers or balanced orientation were classified as “neutral.”

Foundation scores were calculated as the sum of marker weights for each foundation. A post was coded as “using implicit foundations” if it contained any Loyalty, Authority, or Sanctity markers.

3.3.1 Addressing Potential Lexicon-Classification Circularity

A methodological concern arises from the fact that some markers in our lexicon serve dual functions: they contribute both to classifying a post’s political orientation and to identifying its moral foundation content. For instance, “fascista” is coded as a progressive Sanctity marker; a post containing this term is thereby classified as progressive *and* counted as containing a Sanctity marker. This raises the question of whether our finding—that progressive posts contain binding-foundation markers—is an artifact of lexicon construction rather than a genuine empirical result.

We address this concern through three complementary strategies: independent classification validation, theoretical grounding of marker categorization, and analysis of the term’s discursive function.

Independent classification validation. To test whether the lexicon-based classification drives our results, we constructed an alternative classification using hashtags as an independent orientation signal. Posts containing hashtags unambiguously associated with progressive identity (*#sinistra*, *#femminismo*, *#antifascismo*, *#diritti*, *#ambiente*, *#clima*, *#ecologia*) but none associated with populist identity were classified as “tag-progressive,” regardless of their MFT marker content.

This procedure identified 455 posts (Mastodon) classifiable as progressive through hashtags alone that also contained at least one MFT marker. Among these, 85.1% contained binding-foundation markers—virtually identical to the 85.3% obtained through lexicon-based classification. Importantly, the concordance between the two classification methods was only 31.2%, indicating that they capture substantially different subsets of the corpus. The convergence of results despite divergent selection criteria suggests that the finding is robust to classification method.

Theoretical grounding: “Fascista” as contamination language. The classification of “fascista” as a Sanctity marker requires justification independent of our analytical framework. We ground this categorization in two bodies of evidence: historiographical scholarship on the term’s semantic evolution, and corpus-internal analysis of its discursive function.

Historians of fascism have documented the term’s semantic inflation in postwar political discourse. Gentile (2002) warns against “an inflationary use of the category of ‘generic fascism’” (*un uso inflazionistico della categoria di “fascismo generico”*), describing “an ‘elastic’ fascism that continuously expands and contracts in time and space” while “the function of a definition, in the original sense of the term, is to circumscribe, to limit, to fix boundaries.” Orwell (1944), writing as early as 1944, observed that “the word ‘Fascism’ is almost entirely meaningless... by ‘Fascism’ [people] mean, roughly speaking, something cruel, unscrupulous, arrogant, obscurantist, anti-liberal.” Payne (1995) argues that after World War II, “fascism assumed a quasi-religious position within Western culture as a form of absolute moral evil,” giving “the term’s use as an insult a particularly strong form of social power that any other equivalent term lacks.”

This scholarly consensus—that “fascista” functions in contemporary discourse less as a precise ideological descriptor than as a marker of moral exclusion—is consistent with its classification as a Sanctity marker. The term operates according to what Haidt (2012) calls purity/contamination

logic: it marks targets as morally polluted and thereby excludes them from legitimate moral consideration.

Corpus-internal evidence: Target heterogeneity and co-occurrence patterns. Analysis of our corpus confirms the term’s extended, non-referential use. Among 262 Mastodon documents containing "fascis*," target attribution was heterogeneous:

- 5.0% referenced the historical Fascist regime (1922–1945)
- 29.8% referenced contemporary neo-fascist movements or individuals
- 23.7% applied the term to the current Italian government
- 14.9% applied it to the United States or American politics
- 11.5% applied it specifically to Prime Minister Meloni

This distribution indicates that the majority of uses (>60%) extend beyond reference to historically or organizationally fascist entities. The term is applied to mainstream conservative governments, foreign democracies, and individual politicians who do not self-identify as fascist—a pattern consistent with boundary-marking rather than descriptive classification.

Furthermore, "fascista" co-occurs significantly with other exclusionary terms. Table 6 reports that 25.6% of documents containing "fascis*" also contain "nazi*" ($\chi^2 = 66.28, p < .001$ against independence), 9.2% contain "razzis*," and 5.7% contain "regime." This clustering suggests a semantic field organized around moral contamination and political illegitimacy rather than precise ideological classification.

Implications for interpretation. We do not claim that all uses of "fascista" in our corpus are non-referential or that the term never functions as a legitimate ideological descriptor. Some uses clearly reference historical fascism or self-identified neo-fascist movements. Our claim is narrower: the term’s predominant discursive function in progressive Italian social media is to mark moral boundaries—to designate targets as outside the circle of legitimate political consideration. This function is structurally parallel to conservative purity rhetoric that deploys terms like "degenerato" or "perverso" to similar exclusionary effect.

The potential circularity in our methodology is therefore not vicious but reflects a genuine feature of the phenomenon under study: progressive political discourse deploys moral-contamination language, and this language simultaneously marks political orientation and invokes the Sanctity foundation. The two functions are empirically entangled because they are discursively entangled. Our independent validation through hashtag-based classification confirms that this entanglement does not drive the core finding: regardless of how progressive posts are identified, the vast majority contain binding-foundation markers.

3.3.2 Role of the Populist Comparison Group

A referee concern warrants clarification regarding the function of the populist comparison group in our analysis. The populist sample is *not* intended as a control condition in the experimental sense, nor do we claim that it represents a well-defined ideological counterpart to the progressive sample. The Italian political landscape does not map cleanly onto the U.S. liberal-conservative axis that structures most MFT research; Italian “populism” encompasses movements from the nationalist right (Lega, Fratelli d’Italia) to the ideologically heterogeneous Five Star Movement.

The populist sample serves two limited functions:

1. **Lexicon validation:** The fact that our lexicon identifies recognizably different moral vocabularies in posts classified as progressive versus populist provides evidence that the classification captures genuine discursive differences. The contrast confirms that we are not simply detecting generic political language.
2. **Structural comparison:** The between-group analysis reveals how the *same* moral foundations manifest through *different* vocabularies. Populist discourse deploys loyalty through nationalist rhetoric (“patria,” “popolo”); progressive discourse deploys loyalty through movement solidarity. Both invoke the same foundation—the objects and framing differ.

Crucially, the primary hypothesis—that progressive discourse contains binding-foundation markers—is testable without any comparison group. The 85–90% rate of implicit foundation usage among progressive posts is a *within-group* finding that stands independently. Comparison with populist posts enriches the interpretation but does not bear on the core empirical claim. As Ruzza and Fella (2011) note in their analysis of Italian right-wing movements, populist discourse is characterized by “boundary-drawing” and “evocative rhetorical language”—features our lexicon is designed to capture, but which serve an illustrative rather than inferential role in our analysis.

3.4 Statistical Analysis

Given the non-normal distribution of foundation scores (60–90% zero-inflated), we employed both parametric and non-parametric tests to ensure robustness.

Primary tests: Chi-square goodness-of-fit tests assessed whether progressives’ use of implicit foundations exceeded chance levels. Effect size was measured using Cramér’s V .

Between-group comparisons: Mann-Whitney U tests (non-parametric) were used for foundation score comparisons between progressive and populist posts, given the zero-inflated distributions. We report both Cohen’s d (parametric) and Cliff’s δ (non-parametric) effect sizes.

Multiple comparison correction: With 10 between-group tests (5 foundations \times 2 platforms), we applied Holm-Bonferroni correction to control family-wise error rate at $\alpha = .05$.

Effect size interpretation: For Cohen’s d : < 0.2 = negligible, 0.2 – 0.5 = small, 0.5 – 0.8 = medium, > 0.8 = large. For Cliff’s δ : < 0.147 = negligible, 0.147 – 0.33 = small, 0.33 – 0.474 = medium, > 0.474 = large.

4 Results

4.1 Corpus Characteristics

Table 1 presents the corpus characteristics by platform.

Table 1: Corpus Characteristics by Platform

Characteristic	Mastodon	Reddit	Combined
Total documents	2,812	1,977	4,789
With MFT markers	1,239 (44.0%)	1,168 (59.1%)	2,407 (50.3%)
Progressive	682 (24.3%)	636 (32.2%)	1,318 (27.5%)
Populist	402 (14.3%)	389 (19.7%)	791 (16.5%)
Neutral	1,573 (55.9%)	809 (40.9%)	2,382 (49.7%)
Avg. text length	417 chars	910 chars	—

4.2 Main Hypothesis: Implicit Foundation Usage

Table 2 presents the main hypothesis test results. On both platforms, the vast majority of progressive posts contained markers from implicit foundations.

Table 2: Implicit Foundation Usage Among Progressive Posts

Platform	Using Implicit	Total Prog	Rate	χ^2 ($df = 1$)	Cramér's V
Mastodon	582	682	85.3%	340.65***	0.71 (large)
Reddit	573	636	90.1%	408.96***	0.80 (large)
Combined	1,155	1,318	87.6%	—	—

Note: *** $p < .001$. Effect sizes indicate large effects on both platforms.

The breakdown of implicit foundation usage among progressives is shown in Table 3.

Table 3: Implicit Foundation Breakdown (Progressive Posts)

Foundation	Mastodon	Reddit	Combined
Sanctity	413 (39.6%)	474 (36.2%)	887 (37.7%)
Authority	378 (36.2%)	514 (39.3%)	892 (37.9%)
Loyalty	253 (24.2%)	322 (24.6%)	575 (24.4%)

4.3 Foundation Comparison: Progressive vs. Populist

Table 4 presents mean foundation scores by political orientation and platform. Contrary to MFT predictions, progressives scored *higher* than populists on both Sanctity and Authority across both platforms.

Table 4: Mean Foundation Scores by Orientation and Platform

Foundation	Mastodon		Reddit		Direction	
	Prog	Pop	Prog	Pop	Mast	Reddit
Care	0.38	0.29	0.38	0.38	P>C	\approx
Fairness	0.12	0.22	0.19	0.32	C>P	C>P
Loyalty	0.37	0.77	0.51	1.08	C>P	C>P
Authority	0.55	0.43	0.81	0.41	P>C	P>C
Sanctity	0.61	0.44	0.75	0.66	P>C	P>C

Note: P = Progressive, C = Conservative/Populist. Bold indicates results contrary to MFT predictions.

4.4 Statistical Significance

Table 5 presents Mann-Whitney U test results with effect sizes. Following Holm-Bonferroni correction for multiple comparisons, significant differences emerged for Loyalty (both platforms), Authority (both platforms), and Sanctity (Mastodon only).

4.5 Most Frequent Markers

The most frequent markers across both platforms reveal distinct patterns. Progressive discourse was dominated by:

Table 5: Mann-Whitney U Tests with Effect Sizes and Holm-Bonferroni Correction

Foundation	Platform	U	z	p_{corr}	Cohen’s d	Cliff’s δ	Sig.
Care	Mastodon	134,015	0.62	1.000	0.12	0.02	–
	Reddit	121,581	0.46	0.645	–0.01	0.02	–
Fairness	Mastodon	125,418	–2.34	0.077	–0.20	–0.09	–
	Reddit	114,226	–2.06	0.118	–0.20	–0.08	–
Loyalty	Mastodon	101,342	–7.18	<.001	–0.43	–0.26	***
	Reddit	76,908	–10.17	<.001	–0.53	–0.38	***
Authority	Mastodon	114,000	4.64	<.001	0.16	0.17	***
	Reddit	91,240	7.06	<.001	0.40	0.26	***
Sanctity	Mastodon	123,733	2.68	.044	0.18	0.10	*
	Reddit	112,667	2.40	.082	0.09	0.09	–

- **Sanctity contamination labels:** “fascismo” (260 total), “fascista” (166), “razzismo/razzista” (115)
- **Authority appeals:** “europa” (270), “studi” (212), “scienza/esperti” (168)

Populist discourse emphasized:

- **Loyalty markers:** “italiano/i” (300), “nazione” (170), “patria” (123), “popolo” (122)

5 Discussion

5.1 Reinterpreting the Moral Asymmetry

Our cross-platform findings challenge the standard interpretation of MFT’s liberal-conservative asymmetry, though with important nuances revealed by robust statistical analysis.

The primary finding is unambiguous: 85–90% of progressive posts contain binding-foundation markers, with large effect sizes (Cramér’s $V > 0.70$). This directly contradicts the claim that progressives rely exclusively on Care and Fairness. Crucially, this is a *within-group* finding that stands independently of any comparison with other political orientations: the mere presence of binding-foundation markers in the vast majority of progressive posts is sufficient to challenge MFT’s asymmetry thesis.

The between-group comparisons with populist posts serve a different, supplementary function. Rather than establishing that progressives use binding foundations *more* than populists—a claim we do not make—these comparisons illuminate *how* the two groups deploy the same moral-psychological architecture through different vocabularies. The populist group functions as a discursive mirror: it demonstrates that our lexicon successfully captures politically-oriented moral language and reveals the structural parallelism in how both camps invoke loyalty, authority, and purity (cf. Ruzza, 2009, on “exclusionist” discourse patterns).

The breakdown by foundation illustrates this structural parallelism:

- **Authority:** Progressives appeal to science, expertise, and institutional legitimacy; populists appeal to tradition, common sense, and popular sovereignty. Both invoke authority—the objects differ.
- **Sanctity:** Progressives deploy contamination labels targeting ideological deviance (“fascista,” “razzista”); populists deploy contamination labels targeting cultural deviance (“globalista,” “buonista”). Both mark moral pollution—the targets differ.

- **Loyalty:** Populists score significantly higher ($p_{corr} < .001$, medium effect sizes), reflecting explicit nationalist and in-group rhetoric. Progressives show lower but non-negligible loyalty markers, often framed as movement solidarity or coalition-building.

The key insight is that progressives have developed a secularized vocabulary for the binding foundations that allows them to deploy binding-foundation reasoning while maintaining an explicit commitment to individualizing foundations alone. This is not a quantitative claim about who uses more of which foundation, but a structural claim about the shared moral architecture underlying politically divergent discourse.

5.2 "Fascism" as Purity Marker: A Functional Analysis

The term "fascismo/fascista" was the most frequent Sanctity marker in our corpus (426 combined occurrences across platforms). We propose that in contemporary Italian progressive discourse, this term functions primarily as a moral contamination label—a discursive device that marks individuals or ideas as polluted and requiring exclusion from legitimate debate.

To be clear: this analysis does not deny the historical reality of fascism, trivialize its crimes, or suggest that all contemporary uses of the term are analytically invalid. Fascism as a historical phenomenon and as an object of scholarly analysis remains critically important. Our claim is narrower and empirical: in the specific context of social media political discourse, the term "fascista" frequently operates according to the psychological logic of purity/contamination rather than as a precise ideological descriptor.

This interpretation is supported by three observations:

1. **Semantic extension:** The term is applied to a broad range of targets—from explicit neo-fascist movements to mainstream conservative positions, centrist politicians, and even fellow progressives deemed insufficiently committed. This breadth of application suggests a contamination function that exceeds strict definitional boundaries.
2. **Lexical co-occurrence:** "Fascista" frequently co-occurs with other contamination and ideological-boundary terms. Table 6 reports co-occurrence rates for documents containing "fascis*" across both platforms. The term clusters with "nazi" (25.6% Mastodon, 21.4% Reddit), "razzista" (9.2%, 6.8%), "regime" (5.7%, 4.1%), and "odio" (4.2%, 5.5%). A chi-square test confirms that the fascista/nazi co-occurrence significantly exceeds chance expectation ($\chi^2 = 66.28$, $p < .001$), indicating a semantic cluster characteristic of contamination rhetoric rather than independent descriptive usage.
3. **Social function:** Like conservative purity labels ("degenerate," "perverted"), the primary discursive effect is boundary-marking—distinguishing the morally clean in-group from the contaminated out-group.

This functional analysis is consistent with research on metaphorical contamination in moral cognition (Zhong and Liljenquist, 2006; Lee and Schwarz, 2014). The phenomenon also resonates with Italian intellectual debates on the rhetorical function of anti-fascism: Eco (1995) famously identified fourteen features of "Ur-Fascism" precisely because the term resists strict definition, while Galli della Loggia (2018) has critiqued what he calls "il fascismo sempre in agguato e l'antifascismo perenne" (ever-lurking fascism and perpetual anti-fascism) in Italian public discourse. The point is not that progressives are "wrong" to use such language, but that they deploy the same moral-psychological architecture that MFT attributes primarily to conservatives.

5.3 Cross-Platform Robustness

A key strength of this study is the consistency of findings across two distinct platforms:

Table 6: Co-occurrence rates for documents containing "fascis*"

Co-occurring term	Mastodon (%)	Reddit (%)
nazi*	25.6	21.4
razzis*	9.2	6.8
regime	5.7	4.1
odio	4.2	5.5
totalitar*	3.1	1.4
duce	2.3	8.6
ventennio	1.9	3.2

Note: N = 262 (Mastodon), 220 (Reddit) documents containing "fascis*".

Chi-square for fascista/nazi co-occurrence: $\chi^2 = 66.28$, $p < .001$.

- **Mastodon:** Decentralized, shorter posts (avg. 417 chars), more progressive-leaning user base
- **Reddit:** Centralized, longer posts (avg. 910 chars), more diverse user base

Despite these differences, both platforms showed:

- High rates of implicit foundation usage (85-90%)
- Progressives scoring higher than populists on Sanctity
- Progressives scoring higher than populists on Authority
- Populists scoring higher on Loyalty and Fairness

This cross-platform consistency strongly suggests that our findings reflect genuine features of progressive moral discourse rather than platform-specific artifacts.

5.4 Implications for Political Communication

Our findings suggest that progressive and conservative moral discourse may be more similar than MFT implies—both employ all five foundations, but with different vocabularies and explicit self-presentations. This has implications for:

1. **Political persuasion:** Framing progressive positions in terms of Care and Fairness may be less distinctive than commonly assumed
2. **Understanding polarization:** Both sides experience their opponents as morally contaminated, using the same underlying moral logic
3. **Media analysis:** Standard MFT dictionaries may undercount progressive use of binding foundations due to their secularized vocabulary

5.5 Statistical Robustness

Several features of our analysis enhance confidence in the reported findings.

Non-parametric validation. Given the zero-inflated nature of foundation score distributions (60–90% of documents scoring zero on any given foundation), we employed Mann-Whitney U tests rather than t -tests. This non-parametric approach makes no distributional assumptions and is robust to the highly skewed data.

Multiple comparison correction. Applying Holm-Bonferroni correction to the 10 between-group comparisons, we find that Authority differences remain highly significant on both platforms

($p_{corr} < .001$), Loyalty differences remain significant ($p_{corr} < .001$), and Sanctity differences are significant on Mastodon ($p_{corr} = .044$) but not Reddit ($p_{corr} = .082$). Care and Fairness show no significant differences after correction.

Effect sizes. The primary hypothesis (implicit foundation usage) shows large effect sizes (Cramér’s $V = 0.71$ – 0.80). Between-group foundation differences show small to medium effect sizes: Loyalty (Cohen’s $d = 0.43$ – 0.53 , Cliff’s $\delta = 0.26$ – 0.38) and Authority ($d = 0.16$ – 0.40 , $\delta = 0.17$ – 0.26). Sanctity effects are negligible to small ($d = 0.09$ – 0.18 , $\delta = 0.09$ – 0.10).

Interpretation. The main hypothesis—that progressives use implicit foundations—is robustly supported with large effect sizes. The between-group comparisons show that Authority and Loyalty differences are reliable, while Sanctity differences, though consistently in the predicted direction, are modest in magnitude and do not survive correction on Reddit. This nuanced pattern strengthens rather than weakens our theoretical claims: progressives demonstrably use Sanctity rhetoric, but the degree to which they exceed populists is platform-dependent and effect sizes are small.

5.6 Limitations and Scope

Several limitations warrant discussion, though we argue they do not undermine our central findings.

Lexicon-based classification. Our approach relies on a manually constructed lexicon, which inevitably reflects researcher judgment. However, this limitation is shared by all dictionary-based MFT research, including foundational studies (Graham et al., 2009). Our cross-platform replication provides evidence that the findings are not artifacts of idiosyncratic lexicon construction.

Social media as data source. Social media users are not representative of general populations, and online discourse may amplify moral language for performative reasons. However, this limitation applies equally to progressive and populist posts within our corpus. Moreover, if anything, social media’s performative nature should increase alignment with explicit self-presentation—making our finding of implicit binding-foundation usage among progressives more, not less, striking.

Italian specificity. The Italian political context—with its particular history of fascism, strong left-wing tradition, and current populist movements—may not generalize to other national contexts. We view this as an invitation for cross-national replication rather than a fundamental threat to our theoretical contribution. The concept of "secularized purity rhetoric" is designed to be transferable; its specific lexical instantiation will vary across languages and political cultures.

What this study does not claim. We do not argue for moral equivalence between progressive and populist ideologies. We do not deny substantive differences in political goals, policy positions, or empirical accuracy. We do not claim that all uses of "fascist" are analytically inappropriate. Our claim is strictly limited to the moral-cognitive architecture underlying political discourse: progressives and conservatives appear to draw on the same five-foundation structure, differing primarily in vocabulary and explicit self-presentation.

5.7 Future Directions

This study opens several avenues for future research:

1. Cross-linguistic validation with English, French, German, and Spanish discourse
2. Longitudinal analysis of how progressive purity rhetoric has evolved
3. Experimental studies testing whether “fascist” and similar labels trigger disgust responses
4. Survey research comparing explicit endorsement vs. implicit usage of moral foundations

6 Conclusion

Moral Foundations Theory has provided valuable insights into political psychology, but the claim that liberals use only Care and Fairness requires revision. Our cross-platform analysis of 4,789 Italian social media posts reveals that progressives implicitly employ all five moral foundations, with particularly heavy use of Sanctity (purity/contamination) and Authority rhetoric—foundations traditionally associated with conservatism.

The moral vocabularies of progressives and conservatives may differ more in content than in structure: both groups appeal to loyalty, authority, and purity, but target different objects and use different terms. Progressives have developed a secularized purity rhetoric centered on political-ideological contamination labels (“fascist,” “racist,” “toxic”), while maintaining an explicit self-presentation focused on Care and Fairness.

Understanding this hidden moral symmetry may help explain the intensity of contemporary political conflict: both sides experience their opponents as morally contaminated, but neither recognizes that they are using the same underlying moral logic. This suggests that efforts to bridge political divides must address not just the content of disagreements, but the shared moral architecture that structures them.

Acknowledgments

This paper was written with the assistance of Claude Opus 4.5 (Anthropic) and ChatGPT 5.2 (OpenAI), which contributed to data analysis, code development, and manuscript preparation. The author remains solely responsible for all scientific claims and interpretations.

Data Availability

The analysis code (implemented in Pharo Smalltalk) and anonymized datasets are available at www.ai-theoretical.org. Raw data files are provided in both STON (Smalltalk Object Notation) and JSON formats.

References

- Eco, U. (1995). Ur-fascism. In *The New York Review of Books*. Reprinted in Italian as *Il fascismo eterno*, La nave di Teseo, 2018.
- Frimer, J. A., Boghrati, R., Haidt, J., Graham, J., and Dehghani, M. (2019). Moral foundations dictionary 2.0.
- Frimer, J. A., Boghrati, R., Haidt, J., Graham, J., and Dehghani, M. (2020). Do liberals and conservatives use different moral languages? two replications and six extensions of Graham, Haidt, and Nosek’s (2009) moral text analysis. *Journal of Research in Personality*, 84:103906.
- Galli della Loggia, E. (2018). Patologie italiane: il fascismo sempre in agguato e l’antifascismo perenne. In *Speranze d’Italia. Illusioni e realtà nella storia dell’Italia unita*. Il Mulino, Bologna. Originally published 2003.
- Gentile, E. (2002). Fascismo: Storia e interpretazione.
- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., and Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47:55–130.

- Graham, J., Haidt, J., and Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96(5):1029–1046.
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., and Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101(2):366–385.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Vintage Books.
- Haidt, J. and Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4):55–66.
- Haidt, J. and Kesebir, S. (2007). Morality. *Handbook of Social Psychology*, 2:797–832.
- Hoover, J., Portillo-Wightman, G., Yeh, L., Havaladar, S., Davani, A. M., Lin, Y., Kennedy, B., Atari, M., Kazemian, Z., Dehghani, M., et al. (2020). Moral foundations Twitter corpus: A collection of 35k tweets annotated for moral sentiment. *Social Psychological and Personality Science*, 11(8):1057–1071.
- Hopp, F. R., Fisher, J. T., Cornell, D., Huskey, R., and Weber, R. (2021). The extended moral foundations dictionary (eMFD): Development and applications of a crowd-sourced approach to extracting moral intuitions from text. *Behavior Research Methods*, 53:232–246.
- Johnson, K. and Goldwasser, D. (2018). Classification of moral foundations in microblog political discourse. pages 720–730.
- Jost, J. T. and Amodio, D. M. (2009). Political ideology as motivated social cognition: Behavioral and neuroscientific evidence. *Motivation and Emotion*, 33(1):3–23.
- Lai, M., Patti, V., Ruffo, G., and Rosso, P. (2021). The expression of moral values in the Twitter debate: A corpus of conversations. In *IJCoL. Italian Journal of Computational Linguistics*, volume 7, pages 113–132.
- Lee, S. W. and Schwarz, N. (2014). Metaphor in judgment and decision making. pages 657–677.
- Mokhberian, N., Abeliuk, A., Cummings, P., and Lerman, K. (2020). Moral framing and ideological bias of news. pages 206–219.
- Orwell, G. (1944). What is fascism? *Tribune*. Reprinted in *The Collected Essays, Journalism and Letters of George Orwell*, vol. 3.
- Payne, S. G. (1995). *A History of Fascism, 1914–1945*. University of Wisconsin Press, Madison.
- Ruzza, C. (2009). Populism and euroscepticism: Towards uncivil society? *Policy and Society*, 28(1):87–98.
- Ruzza, C. and Fella, S. (2011). Populism and the italian right. *Acta Politica*, 46(2):158–179.
- Shweder, R. A., Much, N. C., Mahapatra, M., and Park, L. (1997). The “big three” of morality (autonomy, community, divinity) and the “big three” explanations of suffering. In *Morality and Health*, pages 119–169. Routledge.
- Turner-Zwinkels, F. M., Johnson, B. B., Sibley, C. G., and Brandt, M. J. (2021). Conservatives’ moral foundations are more densely connected than liberals’ moral foundations. *Personality and Social Psychology Bulletin*, 47(2):167–184.
- Zhong, C.-B. and Liljenquist, K. (2006). Washing away your sins: Threatened morality and physical cleansing. *Science*, 313(5792):1451–1452.