Revista Românească pentru Educație Multidimensională
ISSN: 2066-7329 | e-ISSN: 2067-9270
Abstracting & Indexing | Clarivate Analytics - Web of Science (WoS); EBSCO; Google Scholar; ICI Journals Master List - Index Copernicus; Ideas RePeC; Econpapers; Socionet; CEEOL; Philpapers; KVK; WorldCat; CrossRef; J-GATE

2025, Volume 17, Issue 2, pages: 255-282 | https://doi.org/10.18662/rrem/17.2/982
Submitted: May 19, 2025| *Accepted for publication:* June 4, 2025

Revisiting the Moral Foundations Questionnaire: Integrating Narratives and Bifactor Models

Bogdan POPOVENIUC [1]

[1] „Ştefan cel Mare" University of Suceava, Romania; West University Timişoara, Romania; https://orcid.org/0000-0002-4615-7594; bpopoveniuc@usm.ro

Abstract: *This study examines the validity of the Moral Foundations Questionnaire (MFQ) in the Romanian cultural context, utilizing three independent samples collected over a seven-year period (N=2233). Through multiple analytical approaches, including confirmatory factor analysis, exploratory structural equation modeling, and bifactor modeling, the research reveals significant limitations in existing models of moral psychology, particularly the Moral Foundations Theory (MFT). The analysis of bifactor models suggests that they are better suited to capture the specificity of complex psychological constructs, incorporating significant situational and contextual variability. Building on this, the study proposes a more comprehensive model that integrates the coherence of personal moral narratives and the plasticity of identity into bifactor analysis, offering a deeper understanding of moral judgment. Additionally, the research emphasizes the critical need for exploratory methods and independent replications to continuously enhance the validity and generalizability of assessment instruments. Systemic issues in scientific practices, such as publication bias, lack of replication studies, and pressure for positive results, further obscure the validity of these tools. The study concludes that morality research requires a multidimensional approach that integrates individual narrative processes with social and cultural influences to better capture the complexity of moral phenomena.*

Keywords: *moral constructs, moral identity, self-narrative, integrating factor models, exploratory structural equation modeling.*

How to cite: Popoveniuc, B. (2025). Revisiting the moral foundations questionnaire: integrating narratives and bifactor models. *Revista Românească pentru Educație Multidimensională, 17*(2), 255-282. https://doi.org/10.18662/rrem/17.2/982

©2025 Published by In Manifest Network in partnership with LUMEN Publishing. This is an open access article under the CC BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

Introduction

The Moral Foundations Questionnaire (MFQ), developed by Graham and colleagues (2011), has established itself over the past two decades as one of the most widely used instruments in studying moral phenomena, generating a significant corpus of studies in social psychology. However, in recent years, various studies have questioned its psychometric properties, particularly in cultural contexts different from its original North American setting (Du, 2019). These concerns underscore the importance of continuous validity analysis, not only to ensure reliable and replicable findings but also to highlight opportunities for refining measurement tools and advancing theoretical frameworks.

Building on the MFQ's application in a Romanian cultural context, this study explores its validity through a range of analytical approaches, including confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and bifactor modeling. By critically examining the instrument's strengths and limitations, the study seeks not only to evaluate its theoretical foundations but also to propose innovative methodological improvements. This exploratory approach, akin to reverse epistemological engineering, uses empirical data to refine theoretical models, bridging gaps between measurement tools and real-world moral processes.

In this context, it is proposed that personal (moral) narratives - defined as the stories individuals construct to make sense of their (moral) experiences - constitute a critical variable that can deepen our understanding of moral psychology. The essentialist view of the moral self - grounded in stable traits, as proposed by Moral Foundations Theory (MFT) - is insufficient to fully account for the dynamic, context-sensitive, and dialogical aspects of moral identity. In contrast, the *narrative self*, as proposed by McAdams & Pals (2006), integrates both the fluidity of situational moral meaning-making and the continuity necessary for a coherent moral identity (van den Berg & Corrias, 2023). Recognizing the narrative and dialogical structure of moral selfhood facilitates a more nuanced approach to moral development and expression. Integrating personal moral narratives into psychometric models enhances their sensitivity to individual differences and contextual variability, thus addressing key limitations of static, trait-based assessments. Moreover, methodological tools such as bifactor models - which disentangle general and specific factors - offer a robust framework for capturing the multidimensionality of moral constructs.

Description of the instrument

The MFQ is constructed based on MFT (Graham et al., 2011; Haidt, 2012), which maintains that human morality is organized around five fundamental dimensions. These dimensions are conceived as primary factors of morality and include the following moral intuitions: *care/harm* - which reflects concern for others' welfare and avoidance of suffering, *fairness/cheating* - which expresses sensitivity to justice, equity, and reciprocity, *loyalty/betrayal* - which highlights the importance of in-group attachment and community loyalty, *authority/subversion* - which expresses respect for legitimate social hierarchies and traditions, and *sanctity/degradation* - which highlights concern for physical and spiritual purity, as well as avoidance of behaviors considered dirty or degrading. These five intuitions are further organized into two overarching categories, i.e. Individualizing and Binding. Individualized, autonomous, individual-centered morality encompasses the care/harm and fairness/cheating dimensions and reflects concern for individual rights and welfare. This higher factor is associated with sensitivity to suffering and interpersonal fairness, being more characteristic of those with a more liberal orientation (left-wing in the American context). Group morality ("binding") is centered on group solidarity, includes the loyalty/betrayal, authority/subversion, and sanctity/degradation dimensions, which emphasize group cohesion value, respect for social norms, and spiritual purity. This higher factor is associated with maintaining social order and collective identity and is characteristic of people with conservative and religious orientations.

In constructing the MFQ (Graham et al., 2011, 370), the strategy was to minimize redundancy between items while ensuring comprehensive coverage of the construct. Unlike methods that prioritize achieving high internal consistency, the focus was placed on striking a balance between ensuring internal consistency for the common core and maximizing item heterogeneity to capture to capture comprehensive coverage of the foundation's entire content domain. This approach, the authors considered, allowed capturing various facets of moral foundations.

The final form of the questionnaire includes two scales. A "relevance" subscale designed to cover a broad conceptual area: "When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking?" Specific moral issues that could lead to polarization, such as gun rights or the rights of sexual minorities, are intentionally avoided. Items were designed to reflect various cases of fundamental moral concerns, such as group loyalty in different contexts ("Whether or not someone did something to betray his or her group") or

various types of harm ("Whether or not someone was cruel"). Subjects are asked to evaluate on a scale from 0 - not at all relevant ("This consideration has nothing to do with my judgments of right and wrong") to 5 extremely relevant ("This is one of the most important factors when I judge right and wrong"). A judgement scale was later introduced to counterbalance the subjectivity of personal theory on individual morality (Nisbett & Wilson, 1977) and to mitigate the potential for reduced variance, which could occur if participants consider all aspects to be morally relevant (Graham et al., 2011). This subscale employs multiple response formats to capture genuine moral judgments (e.g., "It is more important to be a team player than to express oneself") rather than relying solely on abstract self-reflections. Participants assess moral scenarios (e.g., "It can never be right to kill a human being") without needing to consciously articulate the underlying basis of their judgments, offering insights into their practical moral reasoning. In this way, the "relevance" subscale involves evaluating explicit theories about what is morally relevant, while the "judgment" subscale evaluates the actual use of moral foundations within judgments. The final version of the MFQ incorporates various item formats to assess sensitivity to fundamental moral concerns, rather than mere opinions about specific issues.

The questionnaire had significant success among researchers exploring topics across various domains of social psychology: political attitudes (Milesi, 2016, 2017; Nilsson et al., 2019; Nilsson & Erlandsson, 2015), religiosity (Sutton et al., 2020), prosocial conduct (Milesi & Alberici, 2018; Nilsson et al., 2016, 2020), attitudes towards refugees/immigrants (Mobayed & Sanders, 2022; Stewart & Morris, 2021), minorities (Argüello-Gutiérrez et al., 2024), and group prejudices (Hadarics & Kende, 2018; Van de Vyver et al., 2016), the poor (Low & Wui, 2016), environmental concerns (Culiberg et al., 2023), or aggressive and adaptive communication traits (Cranmer & Martin, 2015).

However, the validity of the questionnaire appears to be deeply rooted in the cultural framework of American moral-political psychology (Graham et al., 2012), with numerous studies confirming significant and consistent differences in scores across the two higher-order moral orientations between Democrats and Conservative (Blumenau & Lauderdale, 2024; Graham et al., 2009; Hannikainen et al., 2017; Kivikangas et al., 2017; Lewis, 2015; Stewart & Morris, 2021). These differences in the types of intuitions activated during moral judgment provided a natural explanation for the irreconcilable positions of liberals and conservatives on issues such as sexuality and abortion (purity/spirituality), traditional and family values (loyalty and orthodoxy), gun policy, and the tension between individual

rights and state authority (authority and respect). Furthermore, the replication of this structure in an exceptionally large sample - initially tens of thousands and subsequently reaching hundreds of thousands of participants (via yourmorals.org) - along with evidence of structural invariance in studies conducted in other cultural contexts (Davies et al., 2014; Doğruyol et al., 2019; Graham et al., 2009, 2011, 2016) has firmly established this instrument as highly valid and reliable.

However, in recent years, the universality, factorial structure, and validity of the questionnaire (Harper & Rhodes, 2021), and by extension the theory itself, have been increasingly questioned (Tamul et al., 2020). Five-factor models have shown poor data fit, highlighting the need for further refinement (Zakharin & Bates, 2021). An increasing number of studies have raised concerns about internal consistency and the validity of the factorial structure (De Buck & Pauwels, 2023; Harper & Rhodes, 2021; Washburn, 2013; Yalçındağ et al., 2019), particularly in samples that differ in religious beliefs (Davis et al., 2017), racial composition (Davis et al., 2016), or cultural background (Iurino & Saucier, 2020).

For example, Glover et al. (2014) were compelled to remove the relevance items subscale after finding that confirmatory factor analyses failed to produce data-fitting solutions for any reasonable alternative models. Similarly, many studies testing the five-factor model through confirmatory analysis reported only modest model fit (Harper & Rhodes, 2021; Iurino & Saucier, 2020; Kim et al., 2012; Kivikangas et al., 2021; Nilsson & Erlandsson, 2015; Yilmaz et al., 2016), with some even reporting unacceptable levels of fit (Wormley et al., 2023). The fit indices, from most studies that reported them, can be characterized, at best, as "acceptable," often falling below conventional thresholds (e.g., CFI < .90, RMSEA ~ .070) (Bobbio et al., 2011; Cigán et al., 2018; Davies et al., 2014; Glover et al., 2014; Harper & Rhodes, 2021; Iurino & Saucier, 2020; Kim et al., 2012; Kivikangas et al., 2017; Nilsson & Erlandsson, 2015; Wormley et al., 2023; Yilmaz et al., 2016). By contrast, exploratory factor analyses tend to provide stronger support for a two-factor model (Franks & Scherr, 2015; Graham et al., 2011; Johnson et al., 2014; Kugler et al., 2014; Moreira et al., 2019). However, these factors do not always align with the two higher factors - individualizing and binding - predicted by the theory.

Although, the effect size of moral foundations on political choices, as derived from YourMorals.org data, appears to be exaggerated compared to results from independent samples. This discrepancy is likely due, in part, to political interest and self-selection bias - volunteers on the platform being predisposed to have a prior interest and are not representative of the general

population - resulting in inflated effect sizes and a lack of control over repeated questionnaire administrations (Kivikangas et al., 2021).

Additionally, the heterogeneity between the two sets of items used to derive scores for the five foundations - namely, relevance and judgment items, as well as the very status of the five foundations themselves, has increasingly come under scrutiny. Statistical analyses have demonstrated that bifactor models, whether organized around the two types of elements as method factors (i.e., relevance and judgment items) (Wormley et al., 2023) or the two types of morality corresponding to the general moral orientations of Individualizing and Binding (viewed as higher factors in the standard model) (Cigán et al., 2018), exhibit good to very good model fits. For the first model (Wormley et al., 2023), the specificity of the questions in the two sections of the questionnaire - relevance and judgment - operate as two method factors explaining a large part of the response variance, a finding consistent with earlier studies (Akhtar et al., 2023). In the second model (Cigán et al., 2018), the majority of variance is explained by the two higher factors, which basically function as general factors, while the five foundations primarily capture residual variance with limited moral significance. Even in the original study (Graham et al., 2011), the distinction between the two-factor model and the five-factor model was found to be only marginal. This suggests that the five foundations may not stand as independent constructs but instead reflect overlapping or residual components influenced by broader moral dimensions.

After more than a decade of empirical research and numerous unsuccessful attempts at intercultural validation, the authors have introduced a completely revised version of the model featuring six moral foundations. In this updated model, the "Fairness" foundation is split into two distinct foundations: "Equality" and "Proportionality/Equity." In its revised form, the questionnaire includes 36 questions and adopts a different instruction format: "For each of the statements below, please indicate how well each statement describes you or your opinions," inspired by the structure of the more psychometrically robust *Portrait Values Questionnaire* (Schwartz & Bilsky, 1987). Responses are recorded on a 5-point Likert scale: Does not describe me at all (1); Slightly describes me (2); Moderately describes me (3); Describes me fairly well (4); Describes me extremely well (5). This revised format is designed, according to the authors, to better function in cross-cultural contexts (Atari et al., 2023; Matsuo et al., 2019). An increasing number of researchers advocate avoiding the use of the original MFQ, particularly the 20-item version, and instead encourage "to empirically test the psychometric properties of the recently revised MFQ-2, developed by

the authors of MFT as a more accurate instrument for the conceptualization of morality" (De Buck & Pauwels, 2023). This shift is especially critical given that some validity issues have already been identified in replication studies (Winkelkotte et al., 2025).

Another approach to the empirical study and validation of MFT, particularly for understanding the fundamental structure of moral intuitions, involves applying bifactor models to the MFQ. Bifactor models, which assess both general factors and specific factors, have been proposed as a more sophisticated method to capture the complexity of moral foundations compared to traditional single-factor or multifactor models. (Here, the term "bifactor model" refers to models that include at least one general factor orthogonal to the specific factors, while "bifactorial model" will refer to models that feature only two specific factors). This approach is particularly significant given the diverse cultural and ideological contexts in which moral foundations operate, as it offers a more nuanced framework for analyzing the interplay between overarching moral dimensions and specific moral concerns.

The present study is part of a broader research project that spanned multiple cohorts and instruments, initiated seven years ago, at a time when MFT and its measurement instrument were enjoying widespread and enduring popularity in moral psychology. However, as data collection neared completion, the theoretically and empirically validated structure proposed by the authors of the theory appeared increasingly unsupported. In response to growing concerns regarding the instrument's validity, a substantial dataset of over 2,000 participants - divided into three cohorts - was employed to evaluate its measurement properties. This allowed for comprehensive validity analyses based on internal structure and construct validation (Rios & Wells, 2014), providing an opportunity to test theoretically significant models and/or those empirically validated in the specialized literature.

A total of 26 models were tested, encompassing variations with 5, 3, and 2 factors, configured in different ways (e.g., standard, relevance, judgment, with higher factors, bifactor, or hierarchical designs). Analyses focused on reliability, fit indices, and comparative evaluations of standard models.

Data were collected over 6 years, with questionnaires being administered to different student cohorts, the samples generally reflecting biennial chronological order. Given that each cohort had a different relationship to the COVID-19 pandemic: preCOVID-19 (653 participants), preCOVID-19/COVID-19 (436), and COVID-19/postCOVID-19 (1144),

we considered it an opportunity to perform validation analyses as if they were 3 populations with different characteristics.

Reliability analyses

The scale demonstrated good overall reliability across all three samples: preCOVID-19: M = 3.41, SD = .606, α = .884, ω = .894; preCOVID-19/COVID-19: M = 3.38, SD = .571, α = .861, ω = .871; COVID-19/postCOVID-19: M = 3.41, SD = .504, α = .844, ω = .857.

One item, *authority5* ("Men and women each have different roles to play in society"), consistently showed a weak correlation with the rest of the scale (< .2) across all three samples. Confirmatory factor analyses indicated that this item adversely affected model fit across all tested versions, leading to its removal. Notably, across all three samples, this item exhibited a stronger correlation with the Purity foundation, suggesting that, within the Romanian cultural context, gender role attitudes are more closely associated with religiously grounded conceptions of gender (*Purity*) than with hierarchical or institutional aspects typically reflected in the *Authority* foundation.

The reliability indices for most subscales were only marginally acceptable (α > .65), except for the *Purity* scale, which consistently fell within acceptable values. Examination of the correlation matrix revealed that some *judgment* items had weaker correlations (< .3) with their respective subscale items or with the rest of the scale, indicating potential issues with internal consistency.

Comparative analysis of reliability indices for models with 5, 3, and 2 factors shows that reducing the number of factors improves psychometric properties. The 5-factor model presents variable reliability across subscales, with *Purity* showing the strongest reliability and *Authority* the weakest. The 3-factor model has substantially improved reliability, particularly for the *Autonomy* and *Community* factors. The 2-factor model demonstrated the most robust reliability indices for both subscales, *Individualizing* and *Binding*, indicating a more precise and stable measurement of theoretical constructs.

Models fit

The results of the model fitting analyses (CFA) indicate that models with 3 factors and 2 factors tend to have better fit indices across all three samples compared to those with 5 factors, which sometimes fail to converge or produce non-positive definite covariance matrices for latent variables. Across all samples, models with 3 factors (both standard and bifactor

versions) consistently exhibit good fit. Models with 2 factors also yield good results, often with fit indices similar to the 3-factor models.

Another notable finding is that disjoint models based on "relevance" and "judgment" items tend to perform better than combined models, with relevance-based models showing slightly better fit indices compared to judgment-based models. Bifactor-with-3-factors models consistently demonstrate good fit across samples, while bifactor-with-2-factors models also perform well, often achieving fit indices close to the former models. Models with higher factor(s) and 5 factors generally show weaker fit compared to higher factor(s) models with 3 factors, which exhibit reasonable fit but are still outperformed by standard 3-factor or bifactor models.

Regarding fit indices, CFI and TLI values are generally higher for models with 3 and 2 factors, often approaching or exceeding .95. RMSEA values also tend to be lower (indicating better fit) for models with 3 and 2 factors, often below .08, while SRMR values also tend to be lower for models with 3 and 2 factors, frequently below .08.

In conclusion, although the instrument clearly supports the existence of 3 distinct factors (*Autonomy*, *Community*, and *Spirituality*) across all three samples, as predicted by theory, the simple solution with these three correlated factors does not achieve a good fit in any of the samples. Instead, the alternative proposed bifactor models with 3 specific factors - either with 1 general factor, or with 2 distinct general factors for the "relevance" and "judgment" scales (Wormley et al., 2023), or with 2 general factors (*Individualization* and *Group*) (Cigán et al., 2018) - show good or acceptable fit indices.

Table 1 Fit indices for the 3 samples of the bifactor model with 3 specific factors

Bifactor model	χ^2	df	χ^2/df	CFI	TLI	SRMR	RMSEA	RMSEA 90% CI		AGFI	GFI	NFI
								Lower	Upper			
Sample 1	773.547	345	2.242	.979	.975	.056	.053	.048	.058	.972	.982	.962
Sample 2	1129.533	345	3.274	.96	.953	.059	.059	.055	.063	.966	.979	.944
Sample 3	1590.391	345	4.61	.959	.952	.053	.056	.053	.059	.973	.983	.949

The 5-factor model, however, does not meet acceptable fit criteria in any of the samples. These findings underscored the necessity of exploring alternative models, particularly bifactor models, to achieve a more accurate representation of the factorial structure in the Romanian sample.

ESEM-with-CFA analyses

Traditional validation methods, such as exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), have notable limitations. EFA is data-driven and does not impose prior specifications, but the resulting structures can be difficult to compare across groups or over time. On the other hand, CFA imposes predefined relationships between items and factors, neglecting secondary loadings, which compromises the validity of complex models. Post-hoc adjustments aimed at improving CFA models, such as eliminating items or correlating errors, are often criticized for lacking theoretical justification and may negatively impact discriminant validity and subsequent structural analyses (Hu & Bentler, 1998; Morin & Asparouhov, 2018).

In this context, many studies advocate for bifactorial models, which offer robust conceptual justifications, as an alternative to higher-order factor models. A promising option is exploratory structural equation modeling combined with confirmatory factor analysis (ESEM-with-CFA). This approach integrates the flexibility of EFA with the rigor of CFA, enabling the investigation of complex relationships between latent and observed variables (Chen et al., 2006; Howard et al., 2018; Marsh & Alamer, 2024; Marsh et al., 2014; Morin et al., 2020; Swami et al., 2023).

Thus, we evaluated the fit of ESEM-with-CFA models across the three independent samples, testing the three main factorial structures: a 2-factor structure (*Individualzing* and *Binding*), a 3-factor structure (Autonomy, Community, and Spirituality) and a 5-factor structure (*Care*, *Fairness*, *Loyalty*, *Authority*, *Purity*), both with and without a general factor (bifactor). This approach allowed us to assess the empirical support for both theoretical models regarding moral foundations structure.

The analysis of fit indices showed that 5-factor models present superior fit compared to 3 or 2-factor models across all three samples. The 2-factor models demonstrated good fit in the first two samples (CFI = .950-.973, TLI = .942-.969, RMSEA = .059-.066, SRMR = .057-.063), but less satisfactory in the bifactorial version for sample 3 (CFI = .830, TLI = .785).

The 3-factor models showed variable fit indices, from acceptable to good (CFI = .835-.969, TLI = .792-.961, RMSEA = .051-.069, SRMR = .047-.054), with notable improvement in the bifactor version (CFI = .978-.990, TLI = .970-.986, RMSEA = .039-.047).

The 5-factor models demonstrated the best fit across all three samples (CFI = .988-.995, TLI = .982-.993, RMSEA = .028-.037, SRMR

= .04), with additional improvement in the bifactorial version (CFI = .992-.999, TLI = .987-.999, RMSEA = .012-.029, SRMR = .031-.035).

The transition from more parsimonious models with 2 or 3 factors to those with 5 factors led to substantial improvements: ΔCFI between .019 and .16 (the threshold from which it is considered significant being .01), ΔRMSEA between -.02 and -.04 (the significant threshold being -.015), even in the context of more permissive criteria suggested for the WLSMV estimator (Shi et al., 2018; Xia & Yang, 2019). Consistently, the bifactor versions of all models demonstrated superior fit compared to their corresponding standard versions.

Factor loading analysis using ESEM-with-CFA revealed that while the 5-factor model is more stable than the 3-factor model, it exhibits variations in loading consistency across samples, indicating some structural instability.

In the bifactor variant, the general factor (G) shows moderate to strong loadings (>.40) that are generally more consistent across samples for most items, suggesting the presence of a substantial general latent factor underlying the moral foundations. For the simple model, items had stronger loadings on their specific factors, but also showed more significant cross-loadings, reflecting conceptual overlaps between these group-based moral foundations (*Care* and *Fair*) and community-based moral foundations (*Loyalty*, *Authority*, and *Purity*). These overlaps highlight the interconnected nature of these intuitions within their specific moral domain. The 3-factor model exhibited weaker loadings overall, with more pronounced overlaps between specific factors, indicating less distinct representations of moral dimensions. In contrast, the 2-factor model demonstrated better stability, with the *Individualizing* and *Binding* factors being well-represented by moderate to strong loadings. However, cross-loadings and sample-to-sample variations suggest that the moral dimensions are more complex and intertwined than is assumed by theory.

In conclusion, ESEM-with-CFA analysis uncovered a consistent pattern across all three samples, with 5-factor models showing superior fit compared to 3-factor models, contrary to results from CFA solutions, and bifactor versions consistently outperforming their non-bifactor counterparts, offering better representation of the data. While ESEM-with-CFA analysis revealed a factorial structure closely resembling that of MFT, the detected pattern was more complex than the theoretical model. The findings suggest that moral dimensions are not entirely distinct but instead exhibit significant conceptual overlaps, particularly in group-based moral values.

Multi-group invariance

As noted during the presentation of data collection, while the main characteristics of the samples are broadly similar (student), their compositions are heterogeneous in terms of the proportion of specializations, gender, study levels, and their relation to the COVID-19 pandemic period. This variability was considered as an opportunity to validate the model as if the samples represented three distinct populations with differing characteristics, though the results should be interpreted cautiously.

Since sample size significantly affects both the χ^2 fit statistic (Hu & Bentler, 1999; Putnick & Bornstein, 2016) and parameter estimates and fit indices (Kyriazos, 2018), we split the third sample into two equal subsamples through random subjects' extraction. This allowed us to verify model invariance across four samples with 440, 653, 572, and 572 participants, respectively.

Separate analyses of model invariance across the four samples indicated acceptable to a very good fit of the tested models in all cases. All four samples showed acceptable to very good model fit (χ^2/df between 2.358-3.28, CFI between .957-.975, TLI between .950-.971, RMSEA between .055-.059, SRMR=.057-.059). The model fits well for both genders, with minimal differences in fit quality. Configural invariance was supported (χ^2/df=2.732, CFI=.966, RMSEA=.055), confirming the validity of factor structure across all groups. Metric invariance was problematic ($\Delta\chi^2$=879, $p<.001$, ΔCFI=.009, ΔRMSEA<.004), indicating that factor loadings might not be completely equivalent across groups. Scalar invariance was partially supported: $\Delta\chi^2$=337.913, p=.46, suggesting that item intercepts are approximately equal across groups, CFI remains constant (.957 → .957), and RMSEA slightly decreases (.060 → .054), suggesting better model fit. Overall, the results indicate good support for configural invariance, suggesting that the factor structure is consistent across groups. Metric invariance was only borderline supported, indicating variability in factor loadings across groups. Scalar invariance was partially supported, suggesting that item intercepts were roughly equivalent, though not perfectly so.

Discussions

The ESEM-with-CFA analysis clearly supported the 5-factor model (*Care*, *Fairness*, *Loyalty*, *Authority*, and *Purity*), identifying it as superior to the 3-factor model, while highlighting the necessity of the general factor. Conversely, the CFA analysis revealed limitations in the 5-factor model,

showing that it did not fit the data adequately. Instead, the 3-correlated-factors model and the bifactor model with 3 specific correlated factors performed better. Among these, only the bifactor model met acceptable fit criteria, suggesting the presence of a meaningful latent dimension. The stability of this structure was confirmed across all samples. The general factor identified in the bifactor model can be interpreted in several ways. It may represent a general moral dimension overarching the general tendency to attribute importance to moral values, a methodological factor capturing common variation due to response style or social desirability, or a general dimension reflecting specific ontological framing or how individuals perceive and attribute intrinsic moral value to persons and actions (Beal, 2020).

In contemporary psychology, bifactor models have gained increasing ground as valuable tools for understanding complex psychological constructs. Although there have been objections that the superior fit of these models might be attributed to their ability to better adapt to implausible response patterns rather than a more faithful representation of the underlying data structure (Morgan et al., 2015; Reise et al., 2016), numerous recent studies contradict this perspective.

Research conducted by Chen et al. (2016) in personality psychology demonstrated that applying bifactor models not only significantly improves statistical model fit but also results in a notable increase in explained variance for personality factors. This improvement was consistent across all eight samples analyzed and was observed using three different personality measurement instruments, including the Big Five and HEXACO models. The evidence suggests that the general factor identified through these models does not merely represent random error or methodological bias but rather reflects a central psychological mechanism underlying personality structure, associated with processes relevant to individual adaptation.

Similarly, Rodriguez et al. (2016) analyzed indices derived from the standardized loading matrices of bifactor models from 50 published studies. Their findings demonstrated that these indices provide a clearer and more complete understanding of the psychometric properties of assessment instruments. One revealing insight was that, although many measures were described as "multidimensional," the total scores were predominantly influenced by variance attributable to a single latent factor. Additionally, subscale scores often had ambiguous or inconsistent interpretations.

Even major personality constructs, such as the Big Five, remain instrumental approximations, functioning effectively only in the presence of integration with the self's identity narrative (McAdams & Pals, 2006). Chen et al. (2012) further highlighted the utility of bifactor models through their

analysis of the extraversion construct as measured by the NEO-PI-R. They demonstrated that bifactor models are a superior method for simultaneously evaluating such multidimensional constructs as personality. This approach allows for the analysis of both the general construct shared by all facets and the unique specificity of individual facets, independent of the influence of the general construct. A particularly important finding was that, after controlling for the influence of the general construct, extraversion facets exhibited correlations in opposite directions with certain criteria. This underscores the complexity and nuance of the extraversion construct.

In the clinical domain, personality organization levels based on object relations models appear to be better conceptualized in terms of a general factor of personality functioning, as well. Hörz-Sagstetter et al. (2021) found that the general factor explained approximately 66% of the common variance when using the *Inventory of Personality Organization* (IPO). The application of bifactor models has also provided strong support for the concept of the "p" factor in general psychopathology, demonstrating the utility of a trans-diagnostic approach (Caspi et al., 2014). This general factor has proven to be stable over time and effective in describing temporal changes in psychopathology (Gluschkoff et al., 2019). These findings suggest that addressing pathologies of the self cannot be reduced solely to the stable aspects of personality but also requires acknowledgment of the normative component inherent in the concept of self (Glas, 2003).

How can these seemingly contradictory statistical results be explained? In the context of moral psychology, MFT posits that moral intuitions are rapid, automatic, and affective evaluations that guide our moral judgments. While this framework is appealing in its simplicity, it falls short of fully explaining the complexity of moral behavior. To address this, it must be complemented by two key mechanisms: (1) the mechanisms of cognitive processes that trigger the activation of a certain intuitions based on context and individual particularities, and (2) the inclusion of the moral narrative of the self, which provides the interpretation of the events that activates specific intuitions (depending on personal experience, education, and socialization), and also serves as the foundation for behavioral motivation (Krettenauer & Stichter, 2023).

The ADC model (Agent, Deed, Consequences) elucidates the triggering mechanism of moral intuitions by demonstrating that spontaneous moral judgments arise from the preconscious concurrent evaluation of three core elements: the Agent's character/intentions, the Deed (action), and its Consequences (Białek et al., 2014; Dubljević & Racine, 2014). This

framework integrates these components through intuitive processing via cognitive heuristics, advancing Social Intuitionist Theory by specifying the dynamic synthesis of moral valence across each element into a unified judgment. Further grounding this mechanism in the innate deontic structures from universal moral grammar (Mikhail, 2011) may provide complementary foundations for ADC evaluations in moral cognition.

However, the activation of moral intuitions also depends on identification, which functions on two levels: first, situational identification, involving the recognition of the moral nature of a given context; and second, the perceived relationship with the moral identity, referring to the extent to which the situation is seen as relevant to and impacting one's moral self (Glas, 2003). Similar to other personality dimensions, such as narcissism, which shares the common goal of maintaining a positive self-image (Engyel et al., 2022), moral identity represents a dynamic self-regulation process (Blasi, 2005). This process is oriented toward maintaining moral self-image - deontological coherence between moral identity and moral behavior (Arvanitis, 2023; Holyoak & Powell, 2016) - and a morally coherent world (ontological coherence) (Clark et al., 2015). Rather than being a fixed trait, moral identity is better understood as "a complex knowledge structure consisting of moral values, goals, traits, and behavioral scripts" (Aquino et al., 2009, 124). This dynamic structure enables individuals to navigate moral challenges, interpret their experiences, and maintain consistency between their moral self-image and their behavior.

Self-narratives play a crucial role in our moral psychology, operating at multiple levels: personal, interpersonal, social, and cultural (McLean & Syed, 2016). These narratives allow individuals to define and articulate their values, beliefs, and goals, essential for psychological health and resilience (McAdams, 2001). The universality observed in moral foundations (Atari et al., 2023) and value sets (Schwartz & Bilsky, 1987) can be interpreted as reflecting the conceptual structure or fundamental semantic dimensions of the cultures' grand narratives upon which individuals construct the meaning of their social existence. As van den Berg and Corrias (2023, 22) argue: "Given the context specificity of our moral considerations, as well as the richness of our self-narratives in terms of moral meaning, studies should take an exploratory rather than a confirmatory approach to measuring morally relevant phenomena, like moral values." The necessity of bifactor modeling that emerged from the present study - particularly the insights provided by exploratory structural equation modeling analyses - further supports the case for advancing this research methodology in moral psychology.

Limitations and future research directions

The study presents several evident limitations related to the non-representativeness of samples (based exclusively on students and exhibits imbalances in terms of gender and field of study). Furthermore, the inherent weaknesses of the MFQ, common for all questionnaires relies on self-reported data, that is susceptible to biases such as social desirability. The lack of control variables and complementary measurements affects the validity of results, particularly given the influence of the COVID-19 pandemic. Despite these limitations, the findings highlight several promising avenues for future research, including extending and conducting cross-cultural validation of the bifactor solution, exploring the relationship between self-narratives and the general factor, and employing complementary instruments to assess moral identity. Additionally, it would be highly useful integrating these insights with other theoretical models, particularly those investigating the neurocognitive foundations of morality and the interplay between moral intuitions and moral reasoning. By addressing these limitations and pursuing these directions, future studies could provide a more comprehensive understanding of moral psychology, particularly in how values, intuitions, and self-narratives interact to shape moral identity and behavior.

Conclusions

Beyond its specific findings, this research highlights the critical role of integrating quantitative and qualitative methodologies in advancing the study of moral psychology. While the study primarily employed quantitative methods, it underscores the need for mixed-method approaches, particularly in research areas where moral reasoning and identity construction are deeply contextual and narrative-driven. As the dialogical nature of the moral self resists reduction to static traits or dimensions, combining psychometric rigor with narrative inquiry - such as Q methodology (Watts & Stenner, 2012) - is a promising way to enhance our understanding of the complexity and individuality of moral cognition.

The findings also contribute to ongoing debates regarding bifactor models in psychological assessment. While critics argue that improved model fit may arise from accommodating response noise (Bonifay et al., 2017), evidence suggests that bifactor structures effectively capture the interplay between general and specific dimensions of complex constructs (Reise, 2012). In this context, the presence of a general factor of moral judgment - akin to general factors in intelligence (g factor; Johnson et al., 2004), psychopathology (p factor; Smith et al., 2020), or personality (gPF;

Figueredo et al., 2014; Rushton & Irwing, 2009) - invites a reinterpretation: not as an abstract universal trait, but as an expression of the coherence and integration of individual moral narratives. This view supports a more layered and dynamic model of moral identity, calling for both structural and interpretive methodologies.

Furthermore, this study reveals systemic challenges in current research practices. The delayed revision of widely used instruments - such as the MFQ, updated by its own authors after over more than a decade of use - reflects a broader issue: the scarcity of replication studies and institutional incentives for methodological transparency. As only 3% of journals explicitly support submitting replications (Martin & Clarke, 2017), and few platforms encourage the publication of negative or null findings, the reliability and cumulative progress of social science remain vulnerable. Addressing this requires concrete reforms, including pre-registration of studies (Begley & Ellis, 2012), open reporting of all findings, challenges encountered during research, even requiring at least one replication study - whether significant or not - possibly pre-registered (Martin & Clarke, 2017), and the institutionalization of replication efforts. Ultimately, specifically for the meaningful advancement moral psychology, this research emphasizes the importance of integrating personal narratives, multidimensional models, and robust validity checks.

Although this research began seven years ago, the results align with prior findings, confirming the low validity of the MFQ. Our analyses also suggest that the five or six moral foundations proposed by MFT are not redundant, as some radical critics (Gray et al., 2012) have argued. However, they appear to rest on more fundamental psychological processes, such as concern for harm, empathy, cooperation (Curry, 2016, 2019) - an aspect almost completely missing from MFT - or ontological framing, referring to how individuals perceive and attribute intrinsic value to entities within the moral domain (Beal, 2020).

Our statistical findings, along with theoretical and empirical support from other studies, underscore the need for moral psychology to maintain a consistently exploratory dimension, incorporating the coherence and specificity of self-narratives into moral judgment research. Only such an approach better accounts for the flexibility and context-dependence of moral evaluations, shaped by individuals' understandings of the nature of things, rather than rigidly being determined by innate intuitive foundations. Moral identity is constructed and expressed through both evolutionarily and culturally shaped moral intuitions. Social situations elicit affective responses via moral intuitions (Graham et al., 2011), preconsciously mediated by moral heuristics (Dubljević & Racine, 2014), themselves grounded in the generative

structures of a universal moral grammar (Mikhail, 2011). Judgments are conceptually organized via moral schemas (Narvaez & Bock, 2002) or prototypes (Larson, 2013; Rosch, 1973), guided by personal values (Schwartz & Bilsky, 1987), and validated by moral competence (Lind, 2008) – yet constrained by the need for deontological coherence of moral identity. Furthermore, the drive for coherence in moral identity is key to understanding how individuals align their actions and values to maintain a consistent moral worldview. This coherence often blurs the boundary between descriptive and prescriptive judgment, as individuals reshape their descriptive beliefs to support preferred moral conclusions (Clark et al., 2015).

While MFT is well-supported in a broad range of evolutionary, cultural, and anthropological evidence, it nonetheless oversimplifies moral processing by reducing it to automatic intuitive responses based on comparative schemas of social actions, which are subsequently rationalized. In reality, moral judgment involves interpretation from the outset, not merely comparison. It is guided by unconscious heuristics and by personal values formed during culturalization and socialization, influenced by the need for moral coherence (reflected in the general factor), that explains the situational and contextual flexibility of individual moral reasoning.

Acknowledgment

This study is part of the doctoral research project titled *"The Role of Value Orientations and Moral Competence in Explaining Moral Judgments and Decisions,"* conducted at the West University of Timişoara, Romania.

References

Akhtar, N., Francis, L. J., Village, A., Sailer, A. B., Hasan, S. S., & McKenna, U. (2023). Testing the Moral Foundations Questionnaire within a Muslim society: A study among young adults in Pakistan. *Journal of Religious Education*, *71*(1), 1-18. https://doi.org/10.1007/s40839-023-00195-z

Aquino, K., Freeman, D., Reed II, A., Lim, V. K. G., & Felps, W. (2009). Testing a social-cognitive model of moral behavior: The interactive influence of situations and moral identity centrality. *Journal of Personality and Social Psychology*, *97*(1), 123-141. https://doi.org/10.1037/a0015406

Argüello-Gutiérrez, C., López-Rodríguez, L., & Vázquez, A. (2024). The effect of moral foundations on intergroup relations: the salience of fairness promotes the acceptance of minority groups. *Social Psychological and Personality Science*, *15*(1), 93-105. https://doi.org/10.1177/19485506231162161

Arvanitis, A. (2023). Moral self-consistency as the self-organization of moral identity: A social-cognitive approach. *Humanities and Social Sciences Communications*, *10*(1), 1-9. https://doi.org/10.1057/s41599-023-01763-2

Atari, M., Haidt, J., Graham, J., Koleva, S., Stevens, S. T., & Dehghani, M. (2023). Morality beyond the WEIRD: How the nomological network of morality varies across cultures. *Journal of Personality and Social Psychology*, *125*(5), 1157-1188. https://doi.org/10.1037/pspp0000470

Beal, B. (2020). What are the irreducible basic elements of morality? A critique of the debate over monism and pluralism in moral psychology. *Perspectives on Psychological Science*, *15*(2), 273-290. https://doi.org/10.1177/1745691619867106

Begley, C. G., & Ellis, L. M. (2012). Raise standards for preclinical cancer research. *Nature*, *483*(7391), 531-533. https://doi.org/10.1038/483531a

Białek, M., Terbeck, S., & Handley, S. (2014). Cognitive Psychological Support for the ADC Model of Moral Judgment. *AJOB Neuroscience*, *5*(4), 21-23. https://doi.org/10.1080/21507740.2014.951790

Blasi, A. (2005). Moral character: A psychological approach. In D. K. Lapsley & F. C. Power (Eds.), *Character psychology and character education* (pp. 67-100). University of Notre Dame Press.

Blumenau, J., & Lauderdale, B. E. (2024). Liberals and conservatives rely on very similar sets of foundations when comparing moral violations. *American Political Science Review*, 1-17. https://doi.org/10.1017/S0003055424000492

Bobbio, A., Nencini, A., & Sarrica, M. (2011). Il Moral Foundation Questionnaire: Analisi della struttura fattoriale della versione Italiana. [The Moral Foundation Questionnaire. Factorial structure of the Italian version.]. *Giornale Di Psicologia*, *5*(1-2), 7-18. https://hdl.handle.net/11577/2479010

Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science*, *5*(1), 184-186. https://doi.org/10.1177/2167702616657069

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, *2*(2), 119-137. https://doi.org/10.1177/2167702613497473

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*(1), 219-251. https://doi.org/10.1111/j.1467-6494.2011.00739.x

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*(2), 189-225. https://doi.org/10.1207/s15327906mbr4102_5

Chen, Z., Watson, P. J., Biderman, M., & Ghorbani, N. (2016). Investigating the properties of the general factor (M) in bifactor models applied to big five or HEXACO data in terms of method or meaning. *Imagination, Cognition and Personality*, *35*(3), 216-243. https://doi.org/10.1177/0276236615590587

Cigán, J., Kanovský, M., Horský, J., Lane, J. E., & Porubanová, M. (2018). *Five moral foundations as nonmoral residuals. An alternative model for MFQ.* https://www.phil.muni.cz/vyzkum/publikace/prehled/1418916

Clark, C. J., Chen, E. E., & Ditto, P. H. (2015). Moral coherence processes: Constructing culpability and consequences. *Current Opinion in Psychology*, *6*, 123-128. https://doi.org/10.1016/j.copsyc.2015.07.016

Cranmer, G. A., & Martin, M. M. (2015). An examination of aggression and adaption traits with moral foundation. *Communication Research Reports*, *32*(4), 360-366. https://doi.org/10.1080/08824096.2015.1089848

Culiberg, B., Cho, H., Kos Koklic, M., & Zabkar, V. (2023). The role of moral foundations, anticipated guilt and personal responsibility in predicting anti-consumption for environmental reasons. *Journal of Business Ethics: JBE*, *182*(2), 465-481. https://doi.org/10.1007/s10551-021-05016-7

Curry, O. S. (2016). Morality as cooperation: A problem-centred approach. In T. K. Shackelford & R. D. Hansen (Eds.), *The Evolution of Morality* (pp. 27-51). Springer International Publishing. https://doi.org/10.1007/978-3-319-19671-8_2

Curry, O. S. (2019, March 26). What's wrong with moral foundations theory, and how to get moral psychology right. *Behavioral Scientist*. https://behavioralscientist.org/whats-wrong-with-moral-foundations-theory-and-how-to-get-moral-psychology-right/

Davies, C. L., Sibley, C. G., & Liu, J. H. (2014). Confirmatory factor analysis of the Moral Foundations Questionnaire: Independent scale validation in a New Zealand sample. *Social Psychology*, *45*(6), 431-436. https://doi.org/10.1027/1864-9335/a000201

Davis, D. E., Dooley, M. T., Hook, J. N., Choe, E., & McElroy, S. E. (2017). The purity/sanctity subscale of the Moral Foundations Questionnaire does not work similarly for religious versus non-religious individuals. *Psychology of Religion and Spirituality*, *9*(1), 124-130. https://doi.org/10.1037/rel0000057

Davis, D. E., Rice, K., Van Tongeren, D. R., Hook, J. N., DeBlaere, C., Worthington Jr., E. L., & Choe, E. (2016). The moral foundations hypothesis does not replicate well in Black samples. *Journal of Personality and Social Psychology*, *110*(4), e23-e30. https://doi.org/10.1037/pspp0000056

De Buck, A., & Pauwels, L. J. R. (2023). Moral Foundations Questionnaire and Moral Foundations Sacredness Scale: Assessing the Factorial Structure of the Dutch Translations. *Psychologica Belgica*, *63*(1), 92-104. https://doi.org/10.5334/pb.1188

Doğruyol, B., Alper, S., & Yilmaz, O. (2019). The five-factor model of the moral foundations theory is stable across WEIRD and non-WEIRD cultures. *Personality and Individual Differences*, *151*, 109547. https://doi.org/10.1016/j.paid.2019.109547

Du, J. (2019). Validation of the Moral Foundations Questionnaire with three Chinese ethnic groups. *Social Behavior and Personality: An International Journal*, *47*(8), 1-11. https://doi.org/10.2224/sbp.8009

Dubljević, V., & Racine, E. (2014). The ADC of moral judgment: Opening the black box of moral intuitions with heuristics about agents, deeds, and consequences. *AJOB Neuroscience*, *5*(4), 3-20. https://doi.org/10.1080/21507740.2014.939381

Engyel, M., Urbán, R., Bandi, S., & Nagy, L. (2022). Dimensionality of narcissism: A bifactorial model of the Narcissistic Personality Inventory using single-stimulus response formats. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*, *41*(7), 4677-4691. https://doi.org/10.1007/s12144-020-00971-2

Figueredo, A. J., Wolf, P. S. A., Olderbak, S. G., Gladden, P. R., Fernandes, H. B. F., Wenner, C., Hill, D., Andrzejczak, D. J., Sisco, M. M., Jacobs, W. J., Hohman, Z. J., Sefcek, J. A., Kruger, D., Howrigan, D. P., MacDonald, K., & Rushton, J. P. (2014). The psychometric assessment of human life history strategy: A meta-analytic construct validation. *Evolutionary Behavioral Sciences*, *8*(3), 148-185. https://doi.org/10.1037/h0099837

Franks, A. S., & Scherr, K. C. (2015). Using moral foundations to predict voting behavior: Regression models from the 2012 U.S. Presidential election. *Analyses of Social Issues and Public Policy (ASAP)*, *15*(1), 213-232. https://doi.org/10.1111/asap.12074

Glas, G. (2003). Idem, ipse, and loss of the self. *Philosophy, Psychiatry, & Psychology*, *10*(4), 347-352. https://doi.org/10.1353/ppp.2004.0018

Glover, R. J., Natesan, P., Wang, J., Rohr, D., McAfee-Etheridge, L., Booker, D. D., Bishop, J., Lee, D., Kildare, C., & Wu, M. (2014). Moral rationality and intuition: An exploration of relationships between the Defining Issues Test and the Moral Foundations Questionnaire. *Journal of Moral Education*, *43*(4), 395-412. https://doi.org/10.1080/03057240.2014.953043

Gluschkoff, K., Jokela, M., & Rosenström, T. (2019). The general psychopathology factor: structural stability and generalizability to within-individual changes. *Frontiers in Psychiatry*, *10*. https://doi.org/10.3389/fpsyt.2019.00594

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029-1046. https://doi.org/10.1037/a0015141

Graham, J., Meindl, P., Beall, E., Johnson, K. M., & Zhang, L. (2016). Cultural differences in moral judgment and behavior, across and within societies.

Current Opinion in Psychology, *8*, 125-130. https://doi.org/10.1016/j.copsyc.2015.09.007

Graham, J., Nosek, B. A., & Haidt, J. (2012). The moral stereotypes of liberals and conservatives: Exaggeration of differences across the political spectrum. *PloS One*, *7*(12), e50092. https://doi.org/10.1371/journal.pone.0050092

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366-385. https://doi.org/10.1037/a0021847

Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry*, *23*(2), 206-215. https://doi.org/10.1080/1047840X.2012.686247

Hadarics, M., & Kende, A. (2018). Moral foundations of positive and negative intergroup behavior: Moral exclusion fills the gap. *International Journal of Intercultural Relations*, *64*, 67-76. https://doi.org/10.1016/j.ijintrel.2018.03.006

Haidt, J. (2012). *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. Vintage.

Hannikainen, I. R., Miller, R. M., & Cushman, F. A. (2017). Act versus impact: Conservatives and liberals exhibit different structural emphases in moral judgment. *Ratio*, *30*(4), 462-493. https://doi.org/10.1111/rati.12162

Harper, C. A., & Rhodes, D. (2021). Reanalysing the factor structure of the moral foundations questionnaire. *British Journal of Social Psychology*, *60*(4), 1303-1329. https://doi.org/10.1111/bjso.12452

Holyoak, K. J., & Powell, D. (2016). Deontological coherence: A framework for commonsense moral reasoning. *Psychological Bulletin*, *142*(11), 1179-1203. https://doi.org/10.1037/bul0000075

Hörz-Sagstetter, S., Volkert, J., Rentrop, M., Benecke, C., Gremaud-Heitz, D. J., Unterrainer, H.-F., Schauenburg, H., Seidler, D., Buchheim, A., Doering, S., Feil, M. G., Clarkin, J. F., Dammann, G., & Zimmermann, J. (2021). A Bifactor model of personality organization. *Journal of Personality Assessment*, *103*(2), 149-160. https://doi.org/10.1080/00223891.2019.1705463

Howard, J. L., Gagné, M., Morin, A. J. S., & Forest, J. (2018). Using bifactor exploratory structural equation modeling to test for a continuum structure of motivation. *Journal of Management*, *44*(7), 2638-2664. https://doi.org/10.1177/0149206316645653

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424-453. https://doi.org/10.1037/1082-989X.3.4.424

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural*

Equation Modeling: A Multidisciplinary Journal, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Iurino, K., & Saucier, G. (2020). Testing measurement invariance of the moral foundations questionnaire across 27 countries. *Assessment*, *27*(2), 365-372. https://doi.org/10.1177/1073191118817916

Johnson, K. M., Iyer, R., Wojcik, S. P., Vaisey, S., Miles, A., Chu, V., & Graham, J. (2014). Ideology-specific patterns of moral indifference predict intentions not to vote: ideology-specific patterns of moral indifference 1. *Analyses of Social Issues and Public Policy*, *14*(1), 61-77. https://doi.org/10.1111/asap.12039

Johnson, W., Bouchard, T. J., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: Consistent results from three test batteries. *Intelligence*, *32*(1), 95-107. https://doi.org/10.1016/S0160-2896(03)00062-X

Kim, K. R., Kang, J.-S., & Yun, S. (2012). Moral intuitions and political orientation: Similarities and differences between South Korea and the United States. *Psychological Reports*, *111*(1), 173-185. https://doi.org/10.2466/17.09.21.PR0.111.4.173-185

Kivikangas, J. M., Fernández-Castilla, B., Järvelä, S., Ravaja, N., & Lönnqvist, J.-E. (2021). Moral foundations and political orientation: Systematic review and meta-analysis. *Psychological Bulletin*, *147*(1), 55-94. https://doi.org/10.1037/bul0000308

Kivikangas, J. M., Lönnqvist, J.-E., & Ravaja, N. (2017). Relationship of moral foundations to political liberalism-conservatism and left-right orientation in a finnish representative sample. *Social Psychology*, *48*(4), 246-251. https://doi.org/10.1027/1864-9335/a000297

Krettenauer, T., & Stichter, M. (2023). Moral identity and the acquisition of virtue: A self-regulation view. *Review of General Psychology*, *27*(4), 396-409. https://doi.org/10.1177/10892680231170393

Kugler, M., Jost, J. T., & Noorbaloochi, S. (2014). Another look at moral foundations theory: Do authoritarianism and social dominance orientation explain liberal-conservative differences in "moral" intuitions? *Social Justice Research*, *27*(4), 413-431. https://doi.org/10.1007/s11211-014-0223-5

Kyriazos, T. A. (2018). Applied Psychometrics: Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General. *Psychology*, *9*(8), Article 8. https://doi.org/10.4236/psych.2018.98126

Larson, C. (2013). *A Cognitive Prototype Model of Moral Judgment and Disagreement* [Thesis, University of Illinois Chicago]. https://indigo.uic.edu/articles/thesis/A_Cognitive_Prototype_Model_of_Moral_Judgment_and_Disagreement/10949243/1

Lewis, P. (2015). Moral intuitions and smart growth: Why do liberals and conservatives view compact development so differently? *Journal of Urban Affairs*, *37*(2), 87-108. https://doi.org/10.1111/juaf.12172

Lind, G. (2008). The meaning and measurement of moral judgment competence - A dual-aspect model. In D. Fasko, Jr. & W. Willis (Eds.), *Contemporary philosophical and psychological perspectives on moral development and education* (pp. 185-220). Hampton Press.

Low, M., & Wui, Ma. G. L. (2016). Moral foundations and attitudes towards the poor. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*, *35*(4), 650-656. https://doi.org/10.1007/s12144-015-9333-y

Marsh, H., & Alamer, A. (2024). When and how to use set-exploratory structural equation modelling to test structural models: A tutorial using the R package lavaan. *British Journal of Mathematical and Statistical Psychology*, *77*(3), 459-476. https://doi.org/10.1111/bmsp.12336

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: an integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*(1), 85-110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Martin, G. N., & Clarke, R. M. (2017). Are psychology journals anti-replication? a snapshot of editorial practices. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00523

Matsuo, A., Sasahara, K., Taguchi, Y., & Karasawa, M. (2019). Development and validation of the Japanese Moral Foundations Dictionary. *PLoS ONE*, *14*(3), e0213343. https://doi.org/10.1371/journal.pone.0213343

McAdams, D. P. (2001). The psychology of life stories. *Review of General Psychology*, *5*(2), 100-122. https://doi.org/10.1037/1089-2680.5.2.100

McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *The American Psychologist*, *61*(3), 204-217. https://doi.org/10.1037/0003-066X.61.3.204

McLean, K. C., & Syed, M. (2016). Personal, master, and alternative narratives: An integrative framework for understanding identity development in context. *Human Development*, *58*(6), 318-349. https://doi.org/10.1159/000445817

Mikhail, J. (2011). *Elements of moral cognition: Rawls' linguistic analogy and the cognitive science of moral and legal judgment*. Cambridge University Press.

Milesi, P. (2016). Moral foundations and political attitudes: The moderating role of political sophistication. *International Journal of Psychology*, *51*(4), 252-260. https://doi.org/10.1002/ijop.12158

Milesi, P. (2017). Moral Foundations and Voting Intention in Italy. *Europe's Journal of Psychology*, *13*(4), 667-687. https://doi.org/10.5964/ejop.v13i4.1391

Milesi, P., & Alberici, A. I. (2018). Pluralistic morality and collective action: The role of moral foundations. *Group Processes & Intergroup Relations*, *21*(2), 235-256. https://doi.org/10.1177/1368430216675707

Mobayed, T., & Sanders, J. G. (2022). Moral foundational framing and its impact on attitudes and behaviours. *Behavioral Sciences*, *12*(5), Article 5. https://doi.org/10.3390/bs12050118

Moreira, L. V., Souza, M. L. de, & Guerra, V. M. (2019). Validity evidence of a Brazilian version of the Moral Foundations Questionnaire. *Psicologia: Teoria e Pesquisa*, *35*. https://doi.org/10.1590/0102.3772e35513

Morgan, G. B., Hodge, K. J., Wells, K. E., & Watkins, M. W. (2015). Are fit indices biased in favor of bi-factor models in cognitive ability research?: A comparison of fit in correlated factors, higher-order, and bi-factor models via monte carlo simulations. *Journal of Intelligence*, *3*(1), Article 1. https://doi.org/10.3390/jintelligence3010002

Morin, A. J. S., & Asparouhov, T. (2018). *Estimation of a hierarchical Exploratory Structural Equation Model (ESEM) using ESEM-within-CFA*. Substantive Methodological Synergy Research Laboratory. https://www.statmodel.com/download/Webnote%20-%20Hierarchical%20Exploratory%20Structural%20Equation%20Model.pdf

Morin, A. J. S., Myers, N. D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM), and Bifactor-ESEM. In G., Tenenbaum, R. C. Eklund (eds.). *Handbook of sport psychology: Exercise, methodologies, & special topics,* 4th ed (Vol. 2, pp. 1044-1073). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119568124.ch51

Narvaez, D., & Bock, T. (2002). Moral schemas and tacit judgement or how the defining issues test is supported by cognitive science. *Journal of Moral Education*, *31*(3), 297-314. https://doi.org/10.1080/0305724022000008124

Nilsson, A., & Erlandsson, A. (2015). The Moral Foundations taxonomy: Structural validity and relation to political ideology in Sweden. *Personality and Individual Differences*, *76*, 28-32. https://doi.org/10.1016/j.paid.2014.11.049

Nilsson, A., Erlandsson, A., & Västfjäll, D. (2016). The congruency between moral foundations and intentions to donate, self-reported donations, and actual donations to charity. *Journal of Research in Personality*, *65*, 22-29. https://doi.org/10.1016/j.jrp.2016.07.001

Nilsson, A., Erlandsson, A., & Västfjäll, D. (2019). The complex relation between receptivity to pseudo-profound bullshit and political ideology. *Personality & Social Psychology Bulletin*, *45*(10), 1440-1454. https://doi.org/10.1177/0146167219830415

Nilsson, A., Erlandsson, A., & Västfjäll, D. (2020). Moral Foundations Theory and the Psychology of Charitable Giving. *European Journal of Personality*, *34*(3), 431-447. https://doi.org/10.1002/per.2256

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259. https://doi.org/10.1037/0033-295X.84.3.231

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Developmental Review : DR*, *41*, 71-90. https://doi.org/10.1016/j.dr.2016.06.004

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667-696. https://doi.org/10.1080/00273171.2012.715555

Reise, S. P., Kim, D. S., Mansolf, M., & Widaman, K. F. (2016). Is the bifactor model a better model or is it just better at modeling implausible responses? Application of iteratively reweighted least squares to the Rosenberg self-esteem scale. *Multivariate Behavioral Research*, *51*(6), 818-838. https://doi.org/10.1080/00273171.2016.1243461

Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema*, *1*(26), 108-116. https://doi.org/10.7334/psicothema2013.260

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137-150. https://doi.org/10.1037/met0000045

Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, *4*(3), 328-350. https://doi.org/10.1016/0010-0285(73)90017-0

Rushton, J. P., & Irwing, P. (2009). A general factor of personality (GFP) from the multidimensional personality questionnaire. *Personality and Individual Differences*, *47*(6), 571-576. https://doi.org/10.1016/j.paid.2009.05.011

Schwartz, S. H., & Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology*, *53*(3), 550-562. https://doi.org/10.1037/0022-3514.53.3.550

Shi, D., DiStefano, C., McDaniel, H. L., & Jiang, Z. (2018). Examining chi-square test statistics under conditions of large model size and ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(6), 924-945. https://doi.org/10.1080/10705511.2018.1449653

Smith, G. T., Atkinson, E. A., Davis, H. A., Riley, E. N., & Oltmanns, J. R. (2020). The general factor of psychopathology. *Annual Review of Clinical Psychology*, *16*(Volume 16, 2020), 75-98. https://doi.org/10.1146/annurev-clinpsy-071119-115848

Stewart, B. D., & Morris, D. S. M. (2021). Moving morality beyond the in-group: liberals and conservatives show differences on group-framed moral

foundations and these differences mediate the relationships to perceived bias and threat. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.579908

Sutton, G. W., Kelly, H. L., & Huver, M. E. (2020). Political identities, religious identity, and the pattern of moral foundations among conservative Christians. *Journal of Psychology and Theology*, *48*(3), 169-187. https://doi.org/10.1177/0091647119878675

Swami, V., Maïano, C., & Morin, A. J. S. (2023). A guide to exploratory structural equation modeling (ESEM) and bifactor-ESEM in body image research. *Body Image*, *47*, 101641. https://doi.org/10.1016/j.bodyim.2023.101641

Tamul, D., Elson, M., Ivory, J. D., Hotter, J. C., Lanier, M., Wolf, J., & Martínez-Carrillo, N. I. (2020). *Moral Foundations' Methodological Foundations: A Systematic Analysis of Reliability in Research Using the Moral Foundations Questionnaire* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/shcgv

Van de Vyver, J., Houston, D. M., Abrams, D., & Vasiljevic, M. (2016). Boosting Belligerence: How the July 7, 2005, London bombings affected liberals' moral foundations and prejudice. *Psychological Science*, *27*(2), 169-177. https://doi.org/10.1177/0956797615615584

van den Berg, T. G. C., & Corrias, L. D. A. (2023). Moral foundations theory and the narrative self: Towards an improved concept of moral selfhood for the empirical study of morality. *Phenomenology and the Cognitive Sciences*, 1-27. https://doi.org/10.1007/s11097-023-09918-x

Washburn, A. (2013). *Moral Principles and Political Ideology: Exploring the Mediating Role of Abstract Value Endorsements*. College of Science and Health Theses and Dissertations. https://via.library.depaul.edu/csh_etd/53

Watts, S., & Stenner, P. (2012). *Doing Q Methodological Research: Theory, Method & Interpretation*. Sage.

Winkelkotte, F., Fobi, D., Möhring, M., & Wild, S. (2025). Testing psychometrics of the Moral Foundations Questionnaire-2 (MFQ-2) among pre-service teachers in Ghana. *Current Psychology*. *44*,6746-675. https://doi.org/10.1007/s12144-025-07630-4

Wormley, A. S., Scott, M., Grimm, K. J., & Cohen, A. B. (2023). Measuring morality: An examination of the Moral Foundation Questionnaire's factor structure. *Personality and Social Psychology Bulletin*, 01461672231191362. https://doi.org/10.1177/01461672231191362

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods*, *51*(1), 409-428. https://doi.org/10.3758/s13428-018-1055-2

Yalçındağ, B., Türker Özkan, Cesur, S., Yilmaz, O., Tepe, B., Piyale, Z. E., Ali Furkan Biten, & Sunar, D. (2019). An investigation of Moral Foundations

Theory in Turkey using different measures: Research and Reviews. *Current Psychology*, *38*(2), 440-457. https://doi.org/10.1007/s12144-017-9618-4

Yilmaz, O., Harma, M., Bahçekapili, H. G., & Cesur, S. (2016). Validation of the Moral Foundations Questionnaire in Turkey and its relation to cultural schemas of individualism and collectivism. *Personality and Individual Differences*, *99*, 149-154. https://doi.org/10.1016/j.paid.2016.04.090

Zakharin, M., & Bates, T. C. (2021). Remapping the foundations of morality: Well-fitting structural model of the Moral Foundations Questionnaire. *PLoS ONE*, *16*(10), e0258910. https://doi.org/10.1371/journal.pone.0258910