Personality Disorders: Theory, Research, and Treatment

Latent Variable Modeling of Item-Based Factor Scales: Comment on Triarchic or Septarchic?—Uncovering the Triarchic Psychopathy Measure's (TriPM) Structure, by Roy et al.

Christopher J. Patrick, Keanan J. Joyner, Ashley L. Watts, Scott O. Lilienfeld, Antonella Somma, Andrea Fossati, M. Brent Donnellan, Christopher J. Hopwood, Martin Sellbom, Laura E. Drislane, John F. Edens, Edelyn Verona, Robert D. Latzman, Claudio Sica, Stephen D. Benning, Leslie C. Morey, Brian M. Hicks, Kostas A. Fanti, Daniel M. Blonigen, Javier Molto, Mark D. Kramer, and Robert F. Krueger

Online First Publication, October 1, 2020. http://dx.doi.org/10.1037/per0000424

Personality Disorders: Theory, Research, and Treatment

ISSN: 1949-2715

http://dx.doi.org/10.1037/per0000424

COMMENTARY

Latent Variable Modeling of Item-Based Factor Scales: Comment on *Triarchic or Septarchic?—Uncovering the Triarchic Psychopathy Measure's (TriPM) Structure*, by Roy et al.

Christopher J. Patrick and Keanan J. Joyner
Florida State University

Ashley L. Watts
University of Missouri

Scott O. Lilienfeld
Emory University

Antonella Somma and Andrea Fossati
Vita-Salute San Raffaele University

M. Brent Donnellan
Michigan State University

Christopher J. Hopwood
University of California, Davis

Martin Sellbom
University of Otago

Laura E. Drislane
Sam Houston State University

John F. Edens
Texas A&M University

Edelyn Verona
University of South Florida

Robert D. Latzman
Georgia State University

Claudio Sica
University of Florence

Stephen D. Benning
University of Nevada, Las Vegas

Leslie C. Morey
Texas A&M University

Brian M. Hicks
University of Michigan

Kostas A. Fanti
University of Cyprus

Daniel M. Blonigen
Veterans Affairs Palo Alto Health Care System, Livermore, California, and Palo Alto University

Javier Molto
Jaume I University

Mark D. Kramer
Örebro University

Robert F. Krueger
University of Minnesota

Christopher J. Patrick and Keanan J. Joyner, Department of Psychology, Florida State University; Ashley L. Watts, Department of Psychology, University of Missouri; Scott O. Lilienfeld, Department of Psychology, Emory University; Antonella Somma and Andrea Fossati, Department of Psychology, Vita-Salute San Raffaele University; M. Brent Donnellan, Department of Psychology, Michigan State University; Christopher J. Hopwood, Department of Psychology, University of California, Davis; Martin Sellbom, Department of Psychology, University of Otago; Laura E. Drislane, Department of Psychology, Sam Houston State University; John F. Edens, Department of Psychology, Texas A&M University; Edelyn Verona, Department of Psychology, University of South Florida; Robert D. Latzman, Department of Psychology, Georgia State University; Claudio Sica, Department of Psychology, University of Florence; Stephen D. Benning, Department of Psychology, University of Nevada, Las Vegas; Leslie C. Morey, Department of Psychology, Texas A&M University; Brian M. Hicks, Department of Psychiatry, University of Michigan; Kostas A. Fanti, Department of Psychology, University of Cyprus; Daniel M. Blonigen, Veterans Affairs Palo Alto Health Care System, Livermore, California, and Department of Psychology, Palo Alto University; Javier Molto, Department of Basic and Clinical Psychology and Department of Psychobiology, Jaume I University; Mark D. Kramer, Department of Psychology, Örebro University; Robert F. Krueger, Department of Psychology, University of Minnesota.

continued

We critique Roy et al.'s (2020; this issue) approach to characterizing the item-level factor structure of the three scales of the Triarchic Psychopathy Measure (TriPM), in light of the manner in which the TriPM scales were developed, the purposes they were designed to serve, and the growing body of evidence supporting their construct validity. We focus on three major points: (1) The TriPM scales are item-based factor scales – i.e., item sets designed to index broad factors of larger multi-scale (parent) inventories; (2) item-level structural analysis can be useful for representing broad dimensions tapped by such scales, but it cannot be expected to provide an accurate picture of narrower subdimensions (facets) assessed by their parent inventories; and (3) it is critical to consider the nomological networks of the TriPM scales (and other triarchic scale measures) in appraising their effectiveness as operationalizations of the triarchic model constructs. We illustrate the first and second of these points by applying Roy et al.'s analytic approach to the trait scales of the NEO-FFI, which were developed to index broad personality dimensions of the multi-scale NEO-PI-R. We address the third point with reference to the growing body of literature supporting the construct validity of the TriPM scales and demonstrating their utility for advancing an integrative understanding of psychopathy.

Keywords: triarchic model of psychopathy, Triarchic Psychopathy Measure, structural equation modeling, factor analysis, validity

Supplemental materials: http://dx.doi.org/10.1037/per0000424.supp

> Why should we be in such desperate haste to succeed and in such desperate enterprises?
>
> —Thoreau, *Walden* (1906)

Roy et al. (2020) reported item-level factor analyses of the three scales of the Triarchic Psychopathy Measure (TriPM) in six participant samples (*N*s = 150–1,064). They concluded that the "3 original TriPM factors do not optimally represent the conceptual model underlying the TriPM" and that the TriPM in fact indexes "7 emergent factors" (p. 1). We take issue with their analytic approach and conclusions, given the manner in which the TriPM scales were developed, the purposes they were designed to serve, and the growing body of evidence supporting their construct validity.

Our commentary focuses on three points: (a) The TriPM scales are *item-based factor scales*—that is, they comprise items selected to index *broad factors* of larger multiscale (parent) inventories; (b) item-level structural analysis can be useful for *representing broad dimensions* tapped by such scales, but it cannot be expected to provide an accurate picture of narrower subdimensions (facets) assessed by their parent inventories; and (c) it is critical to consider the *nomological networks* of the TriPM scales (and other triarchic scale measures) in appraising their effectiveness as operationalizations of the *triarchic model constructs*. Following description of the triarchic model and TriPM scales, we illustrate the first and second of these points by applying Roy et al.'s analytic approach to the trait scales of the NEO Five-Factor Inventory (NEO-FFI), a short version of the revised NEO Personality Inventory (NEO-PI-R), which were developed to index broad personality dimensions of the multiscale NEO-PI-R. We address the third point in the closing section of this commentary.

Triarchic Model of Psychopathy

The triarchic model was informed in part by the diversity of perspectives and approaches evident in the first edition of Christopher J. Patrick's *Handbook of Psychopathy* (Patrick, 2006). Another impetus for the model was a contentious dispute between prominent investigators in the psychopathy field regarding an article ultimately published in *Psychological Assessment* (Skeem & Cooke, 2010; for details regarding the dispute, see Poythress & Petrila, 2010). This dispute highlighted a longstanding source of acrimony among psychopathy researchers—namely, the failure to distinguish between theoretical constructs and manifest measures.

The triarchic model is a *construct-oriented* model developed to reconcile alternative conceptualizations of psychopathy and assimilate findings across studies using different assessment instruments. It focuses on broad trait dimensions represented in different historic characterizations of psychopathy and measures for assessing it and encourages exploration of their differential external correlates and etiologic bases. As such, the triarchic model is intended to be inclusive and integrative rather than exclusive or sectarian. Other key objectives of the model (Patrick & Drislane, 2015; Patrick, Fowles, & Krueger, 2009) are to (a) allow for *model-based* integration of findings across different studies (see Drislane & Patrick, 2017); (b) facilitate linkages between research on psychopathy in youth and adults through a focus on constructs with clear developmental referents; (c) provide a basis for linking psychopathic symptomatology more effectively with nonreport-based measures by focusing on constructs framed in biobehavioral terms with known physiological correlates; and (d) provide a basis for interfacing findings from psychopathy research with biologi-

Data were provided (in part) by the Human Connectome Project, Washington University in St. Louis and University of Minnesota (WU-Minn) Consortium (principal investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657), funded by the 16 National Institutes of Health Institutes and Centers that support the National Institutes of Health Blueprint for Neuroscience Research and by the McDonnell Center for Systems Neuroscience at Washington University. This article contains no mention of the Personality Inventory for *DSM-5* (PID-5) and no citation or reference to work by Robert F. Krueger.

Correspondence concerning this article should be addressed to Christopher J. Patrick, Department of Psychology, Florida State University, 1107 West Call Street, Tallahassee, FL 32306-4301. E-mail: cpatrick@psy.fsu.edu

cally oriented studies of general psychopathology through a focus on transdiagnostic biobehavioral constructs.

Item-Based Factor Scales

Classic writings on test validity (Cronbach & Meehl, 1955; Messick, 1981) emphasize the *purpose* for which a test is designed and the extent to which a test fulfills its intended purpose. This perspective on validity is reflected within the American Psychological Association's (APA) statements on testing standards (APA, 2014) and ethical principles for assessment (APA, 2017). The *purpose* for which the three scales of the TriPM were created was to provide brief but effective item-based measures of broad factors from two multiscale instruments, the Externalizing Spectrum Inventory (ESI; Krueger, Markon, Patrick, Benning, & Kramer, 2007) and the Boldness Inventory (BI; Patrick et al., 2019)—which can be viewed as the "parent inventories" for the TriPM scales. It bears noting that the TriPM scales, because of their brevity and broad-trait focus, were not intended to effectively index narrower facet dimensions assessed by their parent inventories.

The ESI as a Source of Items for TriPM Meanness and Disinhibition

The ESI (Krueger et al., 2007) was developed to comprehensively assess traits and problems within the externalizing spectrum of psychopathology. A major impetus for its development was behavior genetics research (Krueger et al., 2002), demonstrating a highly heritable factor accounting for appreciable variance in different externalizing disorders and disinhibitory personality traits. Scores on this broad factor correlate with both cognitive-brain measures (Patrick et al., 2006) and cognitive-performance measures (Young et al., 2009), largely as a function of shared genetic influences (Hicks et al., 2007; Young et al., 2009).

The ESI includes *23 unidimensional facet scales* that assess different trait and behavioral expressions of externalizing proneness (general disinhibition; Patrick, Kramer, Krueger, & Markon, 2013). It was constructed using data from both offenders and nonoffenders (overall $N = 1{,}787$) to ensure effective measurement of the full range of each lower order facet dimension as well as the broad externalizing dimension. The ESI includes items with varying levels of difficulty to discriminate effectively at lower and higher trait levels. Follow-up research showed that scores on the general factor of the ESI related in expected ways to diagnostic symptoms and personality traits (Venables & Patrick, 2012), as well as cognitive-brain measures (Nelson, Patrick, & Bernat, 2011; Patrick, Venables, et al., 2013), indicating that its general factor indexes the heritable liability dimension identified by Krueger et al. (2002). As such, it served as one empirical referent (among others) for the *disinhibition construct* of the triarchic model.

Structural analyses of the ESI's 23 subscales also revealed two subsidiary factors, one reflecting callous-aggressive proclivities—defined especially by subscales indexing [lack of] empathy and different expressions of aggression—and the other problematic substance use. Observed correlates of the callous-aggressive subfactor suggested that it tapped characteristics in common with the "affective" features of psychopathy, termed *callous-unemotionality* in the child literature (Frick, Ray, Thornton, & Kahn, 2014). As such, this ESI subfactor served as one empirical referent for the *meanness* construct of the triarchic model.

As reported by Patrick, Kramer, et al. (2013), brief item-based scales were developed to index the ESI's general factor and two subfactors; these scales correspond to the Disinhibition and Meanness scales of the TriPM. The 20-item Disinhibition scale includes items from seven ESI subscales that loaded exclusively onto the ESI's general factor (Krueger et al., 2007, Table 5). The 19-item Meanness scale consists of items from six other ESI subscales that defined the Callous-Aggressive subfactor. Items were selected for each scale based on various considerations including their contribution to effective estimation of the ESI factors, content coverage and nonredundancy, wording polarity, and ability to differentiate disinhibitory tendencies from callous-aggressiveness.

The BI as a Source of Items for TriPM Boldness

The parent inventory for the TriPM Boldness scale, the BI (130 items; Patrick et al., 2019), was developed to index the construct of boldness in terms of distinct but correlated facets using data from a total of 1,791 community participants. The BI's nine unidimensional content scales all load onto a general factor, with some scales loading additionally onto subfactors reflecting emotional stability and venturesomeness. Of note, the BI structural model effectively accommodates the three subscales of the Psychopathic Personality Inventory (PPI/PPI-R; Lilienfeld & Widows, 2005) that define its Fearless Dominance factor—a major empirical referent for the boldness construct of the triarchic model.

The TriPM's 19-item Boldness scale was developed to index the general factor of the BI. Scores on this scale correlate very highly with scores on PPI-Fearless Dominance ($\sim$.8; Patrick et al., 2019; Sellbom & Phillips, 2013) and show parallel relations with criterion measures in domains of physiology, personality, and psychopathology (Patrick, 2018; Patrick & Drislane, 2015; Sellbom, 2018; Sellbom, Laurinavičius, Ustinavičiūtė, & Laurinaitytė, 2018).

Latent Variable Modeling of Item-Based Factor Scales

The TriPM's three scales were designed to index broad latent factors (dimensions) of the ESI and BI, as representations of theoretical constructs of boldness, meanness, and disinhibition. Roy et al.'s analytic strategy neglects the purpose for which these scales were developed, disregards the fact that items within each scale function as indicators of a broad target factor, and focuses on identifying lower order factors defined by subsets of items that cut across distinct ESI facet scales.

To illustrate the problematic aspects of their analytic approach, we report results from work applying this approach to a commonly used instrument whose factor structure is widely accepted, the 60-item NEO-FFI, a short version of the NEO-PI-R (Costa & McCrae, 1992). The NEO-FFI's five scales comprise items from the NEO-PI-R's 30 lower order facets selected to index broad domains of the NEO-PI-R corresponding to the dimensions of the five-factor model of personality (FFM). As such, the NEO-FFI scales are directly analogous to the TriPM scales in terms of their purpose and composition. If Roy et al.'s analytic approach fails to yield the expected five factors for the NEO-FFI, it would raise serious questions regarding its applicability to other broad-trait measures, including the TriPM.

The next section describes samples used in our analyses of the NEO-FFI, steps in the analyses we performed to mirror Roy et al., and results from this approach—then compares these with results obtained using exploratory structural equation modeling (ESEM; see online supplemental materials for additional details). It should be emphasized that this presentation is provided to illustrate our concerns with Roy et al.'s approach by applying it another widely used measure, not to advocate for use of their approach in other contexts.

Problems With Roy et al.'s Analytic Approach: An Empirical Illustration Focusing on the Item Set of the NEO-FFI

Hopwood and Donnellan (2010) critiqued confirmatory factor analysis (CFA) as a technique for modeling the latent structure of multifaceted personality inventories such as the NEO-PI-R, which are rarely marked by simple structure (see also Sellbom & Tellegen, 2019). They instead advocated for an alternative and more psychologically realistic analytic approach—ESEM (see also Marsh et al., 2009; Marsh, Morin, Parker, & Kaur, 2014). ESEM has become the method of choice for modeling the scale-level and item-level structures, respectively, of the NEO-PI-R (Furnham, Guenole, Levine, & Chamorro-Premuzic, 2013; Lui, Samuel, Rollock, Leong, & Chang, 2019) and NEO-FFI (Marsh et al., 2010; Rosellini & Brown, 2011).

Summary of Roy et al.'s Approach

Roy et al. first evaluated the fit of (a) an omnibus CFA model specifying three latent triarchic factors defined by items from the TriPM scales and (b) one-factor CFAs for each individual TriPM scale. Next, they performed separate exploratory factor analyses (EFAs) for the items of each TriPM scale to characterize factors underlying each, and then "removed items with subpar [<.4] loadings . . . and/or that exhibited substantial cross-loadings on other factors that significantly hamper[ed] interpretation of factors" (p. 4). Importantly, removing items that cross-load onto multiple factors *within a single scale* by definition eliminates items that operate to bind the scale together (i.e., that function best as indicators of the general factor underlying various items of the scale). When applied to a scale designed to index a broad construct encompassing different facets, this approach prioritizes item distinctiveness over item coherency and identifies small subsets of items bound together by common elements separate from general-factor variance.

NEO-FFI Data Sets

Our primary analyses used data (N = 1,198; M_{age} = 28.8) from the publicly available Human Connectome Project (HCP; Van Essen et al., 2013); interested readers can access these same data to corroborate our findings or perform their own analyses. We also report results for a separate replication sample (N = 789, M_{age} = 19.4) administered both the NEO-FFI and the TriPM. This research was approved by the institutional review board committee at Florida State University.

Analytic Steps and Findings

Following Roy et al., we used robust weighted least squares estimation to specify CFA/EFA models for the NEO-FFI's ordinal items. Mirroring the fit of the initial omnibus CFA for the TriPM in Roy et al.'s six samples, the correlated five-factor CFA for the full NEO-FFI exhibited inadequate fit in the HCP sample (see Table 1). Notably, the mean loadings of NEO-FFI items on their assigned factors were lower on average than those reported for TriPM items in Roy et al.'s Table 1 ($M\,\lambda$ = .54 for NEO-FFI vs. .65 for TriPM). Also mirroring Roy et al., separate item-level CFAs for the NEO-FFI scales (see Table 1) yielded near-acceptable comparative fit index (CFI) values (M = .88 vs. M = .87 for the TriPM scales) but inadequate values of root mean square error of approximation (RMSEA; M = .11, identical to the TriPM scales).

We then conducted item-level EFAs for the individual NEO-FFI scales and applied Roy et al.'s criterion of ΔCFI >.01 to determine the number of latent factors for each.[1] This indicated a total of 19 factors within the NEO-FFI: three for Neuroticism, five for Extraversion, three for Openness, four for Agreeableness, and four for Conscientiousness. These EFA models fit reasonably well in both the HCP and replication samples (see Table SA in the online supplemental materials), mirroring results for EFAs of the TriPM scales in Roy et al. An omnibus EFA model that extracted 19 factors from the full NEO-FFI item set, reflecting those emerging from the individual-scale EFAs, exhibited acceptable fit (see Table SA in the online supplemental materials). These findings again parallel results for the omnibus seven-factor EFA of the full TriPM item set reported by Roy et al.

Next, we applied Roy et al.'s exclusion criteria—which dropped items with loadings <.40 and/or "substantial" cross-loadings (which we took to mean ≥.40)—to identify items within each scale not robustly or uniquely associated with any particular factor. After dropping items (10 altogether—one each from Neuroticism and Extraversion scales, two each from Openness and Conscientiousness, and four from Agreeableness), three of the 19 factors revealed by the EFAs were defined by only one item. Table 1 shows CFA results for the reduced item sets of the five NEO-FFI scales, excluding single-item factors (two for Extraversion and one for Agreeableness). These models exhibited acceptable fit in the HCP sample, mirroring results for reduced-length TriPM scale CFAs in Roy et al.; acceptable fit was also evident for CFAs of the reduced-length scales in our replication sample (RMSEAs = .02–.08, CFIs = .97–.99, Tucker–Lewis index = .95–.99). As described in the online supplemental material (first subsection under "Results"), the factors "uncovered" for each scale reflected item polarity along with item content (NEO-PI-R facet assignment).

[1] We note that this practice of model comparison is actually inappropriate for multiple reasons. First, Cheung and Rensvold (2002) generated this recommendation in the context of conducting simulation studies comparing measurement invariance models across groups. Second, their simulation studies used maximum likelihood estimation and have not been replicated when using weighted least squares mean- and variance-adjusted estimators (Sass, Schmitt, & Marsh, 2014). And third, it is problematic in general to compare CFI values across models estimated via weighted least squares mean- and variance-adjusted estimators, because the cross-model χ^2 values (which serve as the main formulaic ingredients of CFI) are not on the same scale.

Table 1
Fit Statistics for CFA Models of NEO-FFI Scales in Human Connectome Project Sample (N = 1,198)

NEO-FFI CFA model	CFI	TLI	RMSEA	WLSMV-χ^2(*df*)	*M* loading (range)
Full-original scales					
5-factor omnibus	.74	.73	.07	10,181.85 (1700)	.54 (.16–.80)
1-factor Neuroticism	.93	.92	.09	610.20 (54)	.61 (.46–.83)
1-factor Extraversion	.84	.80	.12	933.38 (54)	.52 (.32–.79)
1-factor Openness	.85	.81	.13	1,187.46 (54)	.50 (.16–.78)
1-factor Agreeableness	.87	.84	.10	693.75 (54)	.53 (.32–.72)
1-factor Conscientiousness	.89	.87	.12	967.21 (54)	.60 (.28–.78)
Reduced-length scales					
16-factor omnibus	.87	.85	.06	4,446.09 (914)	.67 (.30–.89)
3-factor Neuroticism	.97	.96	.07	267.29 (41)	.67 (.55–.80)
3-factor Extraversion	.97	.95	.06	138.12 (24)	.61 (.39–.87)
3-factor Openness	.97	.96	.07	205.41 (32)	.61 (.27–.88)
3-factor Agreeableness	>.99	>.99	.02	15.83 (11)	.70 (.53–.80)
4-factor Conscientiousness	.97	.95	.08	264.21 (29)	.75 (.63–.86)

Note. CFA = confirmatory factor analysis; NEO-FFI = NEO Five-Factor Inventory; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; WLSMV = weighted least squares mean- and variance-adjusted estimators. All negatively keyed items were reverse coded before model fitting to make *M* loading values interpretable. See main text for summary of fit statistics for these models in the replication sample and Table SB in the online supplemental materials for fit statistics of each individual model in the replication sample.

Compared with the initial five-factor CFA model, fit improved for an omnibus 16-factor model of the reduced NEO-FFI item set (i.e., 47 items across five scales, excluding ones indicative of single-item factors) and was comparable with that of the seven-factor omnibus TriPM model across samples in Roy et al. (*M* CFI = .87 vs. .89, *M* RMSEA = .05 vs. .05; see Table 1 and Table SA in the online supplemental materials).

Following Roy et al.'s logic, one might conclude from these analyses—substituting "NEO-FFI" for "TriPM" and "FFM" for "triarchic model"—that (a) "the [NEO-FFI] may not accord sufficiently with the [FFM of personality]" (p. 7), (b) "the [FFM] domains cannot be represented via an item-level [FFM], given [that] each [NEO-FFI] scale is clearly multidimensional" (p. 7–8), (c) "the [five] original [NEO-FFI] scales are misrepresenting important sources of covariation, and therefore, the [five-factor] model is mis-specified (i.e., does not accurately account for the structure of [NEO-FFI] item covariance" (p. 8)), (d) "results support use of the [sixteen] first-order factors as a guide for forming new composites with the [NEO-FFI] items" (p. 11), and (e) "continued use of five original [NEO-FFI] scales can lead to theoretical ambiguity and statistical washout effects, which will hinder our understanding of [personality constructs] with huge impact" (p. 13).

We object to these conclusions—whether applied to the NEO-FFI or the TriPM. For each inventory, there is a strong a priori expectation that its scales are each undergirded by a dominant broad factor because items of each scale were selected to index a general factor encompassing distinct facets represented in parent inventories. The presence of a dominant factor within each scale is evident from their scree plots, which in each case reveal a first eigenvalue exceeding the second by three or more times (Morizot, Ainsworth, & Reise, 2007). The presence of additional factors in each scale reflects diversity in thematic content and item characteristics such as wording, polarity, and difficulty that contribute to covariation patterns (see first subsection of online supplemental results). Importantly, the general factors indexed by the five scales of the NEO-FFI—corresponding to dimensions of the FFM as indexed by the NEO-PI-R domains—have well-developed nomological networks that need to be considered in appraising their validity (Hopwood & Donnellan, 2010; Morey, 2019). We return to the important issue of external validity evidence for construct operationalizations in the final section of this commentary.

Comparative Results Using ESEM

Given these considerations, an alternative to the analytic approach of Roy et al. is needed to effectively model the broad latent factor underlying each of the NEO-FFI scales—one that does not rely on unrealistic expectations of simple structure and that accounts for covariation among particular subsets of items apart from their shared associations with the general factor. Somma, Borroni, Drislane, Patrick, and Fossati (2019) reported an alternative, target-rotated ESEM approach for this purpose, which they applied to the item-based factor scales of the TriPM. Their ESEM approach modeled a general factor for each scale while accounting for covariation among particular subsets of items not captured by the general factor and yielded good fit for an omnibus three-factor model incorporating all 58 items of the TriPM.

We briefly summarize results from counterpart ESEM models of the NEO-FFI here; further details regarding the rationale for ESEM, the specific analytic procedures of Somma et al. (2019), and findings from analyses applying these procedures to the NEO-FFI are provided in the online supplemental materials. As shown in Table 2, the five-factor target rotated ESEM for the full NEO-FFI item set exhibited excellent fit in both the HCP and replication samples. Speaking to Roy et al.'s concern about correlated residual terms, the model fit somewhat less well in these two samples without inclusion of correlated residuals (*M* CFI = .90, *M* RMSEA = .04)—but fit remained comparable with that of Roy et al.'s final seven-factor TriPM model (corresponding *M*s = .89 and .05).

Given that TriPM data were also available for the replication sample, we modeled the full TriPM item set using Somma et al.'s ESEM approach and found good fit (as did Latzman et al., 2019 and Paiva et al., 2020) for an omnibus three-factor model—both

Table 2
Fit Statistics for Five-Factor Omnibus ESEM With Target Rotation of Full-Original NEO-FFI Scales in Each Participant Sample

Participant sample	CFI	TLI	RMSEA	WLSMV-χ^2(*df*)
HCP sample (*N* = 1,198)	.94	.92	.03	3,430.20 (1416)
Replication sample (*N* = 789)	.94	.93	.03	2,421.48 (1416)

Note. ESEM = exploratory structural equation model; NEO-FFI = NEO Five-Factor Inventory; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; WLSMV = weighted least squares mean- and variance-adjusted estimators; HCP = Human Connectome Project.

with and without inclusion of correlated residuals (CFIs = .99/.98, Tucker–Lewis index = .99/.98, RMSEAs = .04/.05; see online supplemental material for further details). We encourage Roy et al. to report findings for this omnibus ESEM model, with and without correlated residuals, in their participant samples.[2]

Conclusions

The foregoing results indicate that ESEM provides an effective method for modeling broad factors corresponding to N, E, O, A, and C using all items of the NEO-FFI, and broad factors corresponding to boldness, meanness, and disinhibition using all items of the TriPM. By contrast, Roy et al.'s method failed to model the broad dimensions these scales were designed to measure, and instead delineated narrow subfactors with poor resemblance to known facets of these broad dimensions. Lower order facets of N, E, O, A, and C could not be recovered (or "uncovered") from the items of the NEO-FFI owing to their incomplete representation in these brief-form scales. Accordingly, researchers wishing to examine associations for lower order facets of the FFM are encouraged to use the NEO-PI-R (or its successor, the NEO-PI-3; [McCrae & Costa, 2010]), which provides effective facet-level coverage. Similarly, researchers wishing to examine associations for lower order facets of boldness, meanness, and disinhibition are encouraged to use relevant facet scales of the ESI and the BI.

Nomological Networks and Construct Validity

As a final point, broad factors of inventories such as the NEO-PI-R, the ESI, and the BI gain meaning and demonstrate utility through growing knowledge of their nomological networks—that is, their convergent and discriminant associations with measures of other types. Although investigations of the internal structure of measures are important, investigations of what Loevinger (1957) termed *external validity* are equally if not more critical when appraising their construct validity (Hopwood & Donnellan, 2010; Lykken, 1971).[3]

The NEO Personality Inventory, as a widely used FFM measure that has been available for many years, has a well-developed nomological network and has proven useful for many important purposes. Although the nomological network of the newer TriPM is less well-developed, considerable progress has been made in this regard (see Sellbom, 2018), and it has proven useful as a referent for comparing the content coverage of different psychopathy inventories (Drislane, Patrick, & Arsal, 2014; Sellbom & Phillips, 2013), for bridging the youth and adult psychopathy literatures (Sica, Ciucci, Baroncelli, Frick, & Patrick, 2019; Somma, Borroni, Drislane, & Fossati, 2016), and as a clinical-predictive tool (Sellbom, Lilienfeld, Fowler, & McCrary, 2018). In addition, the TriPM (and other scales developed to index the triarchic model constructs; see, e.g.: Brislin et al., 2019; Drislane & Patrick, 2017) has been helpful for understanding associations of different physiological measures including aversive startle potentiation, oddball-P3 brain response, and fear–face brain reactivity with psychopathy (Patrick, 2018) and other clinical phenomena (Patrick, Venables, et al., 2013; Venables et al., 2017, 2018; Yancey, Venables, & Patrick, 2016).

From this standpoint, we view the TriPM and other triarchic-scale measures as serving aims complementary to those of the FFM and psychopathy inventories such as the Psychopathy Checklist-Revised (Hare, 2003) and the Self-Report Psychopathy scale (Paulhus, Neumann, & Hare, 2017). We believe that using and comparing different conceptualizations and measurement methods is valuable for advancing the study of psychopathy by allowing for integration and synthesis of the best supported features of each. Although we recognize that others may view alternative approaches in more competitive terms, it is important for all researchers to be cognizant of the purposes for which particular measures are designed, mindful of their nomological networks, and willing to apply the same standards regardless of personal scholarly preferences when selecting statistical modeling methods to apply to them.

[2] Our point regarding the appropriateness of an ESEM approach to modeling the structure of item-based factor scales applies also to the work of Collison, Miller, and Lynam (2020), who relied exclusively on EFA with parallel analysis to characterize the item-level structure of alternative triarchic (Tri-) scale measures and compare their content coverage. We encourage these and other investigators to model general factors of alternative Tri-scale measures using ESEM, as basis for evaluating their relative coherency and convergent/discriminant relations with one another.

[3] This point regarding the importance of external criterion correlations as a basis for evaluating scale validity applies also to the work of Collison et al. (2020).

References

American Psychological Association. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16,* 397–438. http://dx.doi.org/10.1080/10705510903008204

Brislin, S. J., Patrick, C. J., Flor, H., Nees, F., Heinrich, A., Drislane, L. E., . . . Foell, J. (2019). Extending the construct network of trait disinhibition to the neuroimaging domain: Validation of a Bridging Scale for Use in the European IMAGEN project. *Assessment, 26,* 567–581. http://dx.doi.org/10.1177/1073191118759748

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5

Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment, 31,* 1412–1427. http://dx.doi.org/10.1037/pas0000626

Collison, K. L., Miller, J. D., & Lynam, D. R. (2020). Examining the factor structure and validity of the Triarchic Model of Psychopathy across measures. *Personality Disorders: Theory, Research, and Treatment*. Advance online publication. http://dx.doi.org/10.1037/per0000394

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI–R) and the NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302. http://dx.doi.org/10.1037/h0040957

Drislane, L. E., & Patrick, C. J. (2017). Integrating alternative conceptions of psychopathic personality: A latent variable model of triarchic psychopathy constructs. *Journal of Personality Disorders, 31,* 110–132. http://dx.doi.org/10.1521/pedi_2016_30_240

Drislane, L. E., Patrick, C. J., & Arsal, G. (2014). Clarifying the content coverage of differing psychopathy inventories through reference to the triarchic psychopathy measure. *Psychological Assessment, 26,* 350–362. http://dx.doi.org/10.1037/a0035152

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin, 140,* 1–57. http://dx.doi.org/10.1037/a0033076

Furnham, A., Guenole, N., Levine, S. Z., & Chamorro-Premuzic, T. (2013). The NEO Personality Inventory-Revised: Factor structure and gender invariance from exploratory structural equation modeling analyses in a high-stakes setting. *Assessment, 20,* 14–23. http://dx.doi.org/10.1177/1073191112448213

Hare, R. D. (2003). *The Hare Psychopathy Checklist–Revised* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Hicks, B. M., Bernat, E., Malone, S. M., Iacono, W. G., Patrick, C. J., Krueger, R. F., & McGue, M. (2007). Genes mediate the association between P3 amplitude and externalizing disorders. *Psychophysiology, 44,* 98–105. http://dx.doi.org/10.1111/j.1469-8986.2006.00471.x

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review, 14,* 332–346. http://dx.doi.org/10.1177/1088868310361240

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179–185. http://dx.doi.org/10.1007/BF02289447

Krueger, R. F., Hicks, B. M., Patrick, C. J., Carlson, S. R., Iacono, W. G., & McGue, M. (2002). Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum. *Journal of Abnormal Psychology, 111,* 411–424. http://dx.doi.org/10.1037/0021-843X.111.3.411

Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D., & Kramer, M. D. (2007). Linking antisocial behavior, substance use, and personality: An integrative quantitative model of the adult externalizing spectrum. *Journal of Abnormal Psychology, 116,* 645–666. http://dx.doi.org/10.1037/0021-843X.116.4.645

Latzman, R. D., Palumbo, I. M., Sauvigné, K. C., Hecht, L. K., Lilienfeld, S. O., & Patrick, C. J. (2019). Psychopathy and internalizing psychopathology: A triarchic model perspective. *Journal of Personality Disorders, 33,* 262–287. http://dx.doi.org/10.1521/pedi_2018_32_347

Lilienfeld, S. O., & Widows, M. (2005). *Psychopathic Personality Inventory–Revised professional manual*. Odessa, FL: PAR.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3,* 635–694. http://dx.doi.org/10.2466/pr0.1957.3.3.635

Lui, P. P., Samuel, D. B., Rollock, D., Leong, F. T. L., & Chang, E. C. (2019). Measurement invariance of the five factor model of personality: Facet-level analyses among Euro and Asian Americans. *Assessment, 27,* 887–902. http://dx.doi.org/10.1177/1073191119873978

Lykken, D. T. (1971). Multiple factor analysis and personality research. *Journal of Experimental Research in Personality, 5,* 161–170.

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22,* 471–491. http://dx.doi.org/10.1037/a0019227

Marsh, H. W., Morin, A. J., 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,* 85–110. http://dx.doi.org/10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling, 16,* 439–476. http://dx.doi.org/10.1080/10705510903008220

McCrae, R. R., & Costa, P. T. (2010). *NEO inventories for the NEO Personality Inventory-3 (NEO-PI-3), NEO Five-Factor Inventory-3 (NEO-FFI-3), and NEO Personality Inventory-Revised (NEO PI-R): Professional manual*. Lutz, FL: PAR.

Messick, S. (1981). Evidence and ethics in the evaluation of tests. *Educational Researcher, 10,* 9–20. http://dx.doi.org/10.3102/0013189X010009009

Morey, L. C. (2019). Thoughts on the assessment of the *DSM–5* alternative model for personality disorders: Comment on Sleep et al. *Psychological Assessment, 31,* 1192–1199. http://dx.doi.org/10.1037/pas0000710

Morizot, J., Ainsworth, A. T., & Reise, S. P. (2007). Toward modern psychometrics: Application of item response theory models in personality research. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 407–423). New York, NY: Guilford Press.

Nelson, L. D., Patrick, C. J., & Bernat, E. M. (2011). Operationalizing proneness to externalizing psychopathology as a multivariate psychophysiological phenotype. *Psychophysiology, 48,* 64–72. http://dx.doi.org/10.1111/j.1469-8986.2010.01047.x

Paiva, T. O., Pasion, R., Patrick, C. J., Moreira, D., Almeida, P. R., & Barbosa, F. (2020). Further evaluation of the Triarchic Psychopathy Measure: Evidence from community adult and prisoner samples from Portugal. *Psychological Assessment, 32,* e1–e14. http://dx.doi.org/10.1037/pas0000797

Patrick, C. J. (2006). *Handbook of psychopathy*. New York, NY: Guilford Press.

Patrick, C. J. (2018). Cognitive and emotional processing in psychopathy. In C. J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed., pp. 422–455). New York, NY: Guilford Press.

Patrick, C. J., Bernat, E. M., Malone, S. M., Iacono, W. G., Krueger, R. F., & McGue, M. (2006). P300 amplitude as an indicator of externalizing in adolescent males. *Psychophysiology, 43,* 84–92. http://dx.doi.org/10.1111/j.1469-8986.2006.00376.x

Patrick, C. J., & Drislane, L. E. (2015). Triarchic model of psychopathy: Origins, operationalizations, and observed linkages with personality and general psychopathology. *Journal of Personality, 83,* 627–643. http://dx.doi.org/10.1111/jopy.12119

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology, 21,* 913–938. http://dx.doi.org/10.1017/S0954579409000492

Patrick, C. J., Kramer, M. D., Krueger, R. F., & Markon, K. E. (2013). Optimizing efficiency of psychopathology assessment through quantitative modeling: Development of a brief form of the Externalizing Spectrum Inventory. *Psychological Assessment, 25,* 1332–1348. http://dx.doi.org/10.1037/a0034864

Patrick, C. J., Kramer, M. D., Vaidyanathan, U., Benning, S. D., Hicks, B. M., & Lilienfeld, S. O. (2019). Formulation of a measurement model for the boldness construct of psychopathy. *Psychological Assessment, 31,* 643–659. http://dx.doi.org/10.1037/pas0000690

Patrick, C. J., Venables, N. C., Yancey, J. R., Hicks, B. M., Nelson, L. D., & Kramer, M. D. (2013). A construct-network approach to bridging diagnostic and physiological domains: Application to assessment of externalizing psychopathology. *Journal of Abnormal Psychology, 122,* 902–916. http://dx.doi.org/10.1037/a0032807

Paulhus, D. L., Neumann, C. S., & Hare, R. D. (2017). *Manual for the Self-Report Psychopathy Scale* (4th ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Poythress, N., & Petrila, J. P. (2010). PCL-R Psychopathy: Threats to sue, peer review, and potential implications for science and law: A commentary. *The International Journal of Forensic Mental Health, 9,* 3–10. http://dx.doi.org/10.1080/14999013.2010.483346

Rosellini, A. J., & Brown, T. A. (2011). The NEO Five-Factor Inventory: Latent structure and relationships with dimensions of anxiety and depressive disorders in a large clinical sample. *Assessment, 18,* 27–38. http://dx.doi.org/10.1177/1073191110382848

Roy, S., Vize, C., Uzieblo, K., van Dongen, J. D. M., Miller, J., Lynam, D., . . . Neumann, C. S. (2020). Triarchic or septarchic?-Uncovering the triarchic psychopathy measure's (TriPM) structure. *Personality Disorders: Theory, Research, and Treatment*. Advance online publication. http://dx.doi.org/10.1037/per0000392

Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling, 21,* 167–180. http://dx.doi.org/10.1080/10705511.2014.882658

Sellbom, M. (2018). The triarchic psychopathy model: Theory and measurement. In M. DeLisi (Ed.), *Routledge international handbook of psychopathy and crime* (pp. 241–264). New York, NY: Routledge. http://dx.doi.org/10.4324/9781315111476-16

Sellbom, M., Laurinavičius, A., Ustinavičiūtė, L., & Laurinaitytė, I. (2018). The Triarchic Psychopathy Measure: An examination in a Lithuanian inmate sample. *Psychological Assessment, 30,* e10–e20. http://dx.doi.org/10.1037/pas0000603

Sellbom, M., Lilienfeld, S. O., Fowler, K. A., & McCrary, K. L. (2018). The self-report assessment of psychopathy: Challenges, pitfalls, and promises. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 211–258). New York, NY: Guilford Press.

Sellbom, M., & Phillips, T. R. (2013). An examination of the triarchic conceptualization of psychopathy in incarcerated and nonincarcerated samples. *Journal of Abnormal Psychology, 122,* 208–214. http://dx.doi.org/10.1037/a0029306

Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment, 31,* 1428–1441. http://dx.doi.org/10.1037/pas0000623

Sica, C., Ciucci, E., Baroncelli, A., Frick, P. J., & Patrick, C. J. (2019). Not just for adults: Using the triarchic model of psychopathy to inform developmental models of conduct problems in adolescence. *Journal of Clinical Child and Adolescent Psychology*. Advance online publication. http://dx.doi.org/10.1080/15374416.2019.1574228

Skeem, J. L., & Cooke, D. J. (2010). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment, 22,* 433–445. http://dx.doi.org/10.1037/a0008512

Somma, A., Borroni, S., Drislane, L. E., & Fossati, A. (2016). Assessing the triarchic model of psychopathy in adolescence: Reliability and validity of the Triarchic Psychopathy Measure (TriPM) in three samples of Italian community-dwelling adolescents. *Psychological Assessment, 28,* e36–e48. http://dx.doi.org/10.1037/pas0000184

Somma, A., Borroni, S., Drislane, L. E., Patrick, C. J., & Fossati, A. (2019). Modeling the structure of the triarchic psychopathy measure: Conceptual, empirical, and analytic considerations. *Journal of Personality Disorders, 33,* 470–496. http://dx.doi.org/10.1521/pedi_2018_32_354

Thoreau, H. D. (1906). *The writings of Henry David Thoreau: II. Walden*. Boston, MA: Houghton Mifflin & Co.

Turkheimer, E., Ford, D. C., & Oltmanns, T. F. (2008). Regional analysis of self-reported personality disorder criteria. *Journal of Personality, 76,* 1587–1622. http://dx.doi.org/10.1111/j.1467-6494.2008.00532.x

van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & Wu-Minn, H. C. P. (2013). The WU-Minn human connectome project: An overview. *NeuroImage, 80,* 62–79. http://dx.doi.org/10.1016/j.neuroimage.2013.05.041

Venables, N. C., Foell, J., Yancey, J. R., Kane, M. J., Engle, R. W., & Patrick, C. J. (2018). Quantifying inhibitory control as externalizing proneness: A cross-domain model. *Clinical Psychological Science, 6,* 561–580. http://dx.doi.org/10.1177/2167702618757690

Venables, N. C., & Patrick, C. J. (2012). Validity of the Externalizing Spectrum Inventory in a criminal offender sample: Relations with disinhibitory psychopathology, personality, and psychopathic features. *Psychological Assessment, 24,* 88–100. http://dx.doi.org/10.1037/a0024703

Venables, N. C., Yancey, J. R., Kramer, M. D., Hicks, B. M., Nelson, L. D., Strickland, C. M., . . . Patrick, C. J. (2017). Evidence of a prominent genetic basis for associations between psychoneurometric traits and common mental disorders. *International Journal of Psychophysiology, 115,* 4–12. http://dx.doi.org/10.1016/j.ijpsycho.2016.09.011

Yancey, J. R., Venables, N. C., & Patrick, C. J. (2016). Psychoneurometric operationalization of threat sensitivity: Relations with clinical symptom and physiological response criteria. *Psychophysiology, 53,* 393–405. http://dx.doi.org/10.1111/psyp.12512

Young, S. E., Friedman, N. P., Miyake, A., Willcutt, E. G., Corley, R. P., Haberstick, B. C., & Hewitt, J. K. (2009). Behavioral disinhibition: Liability for externalizing spectrum disorders and its genetic and environmental relation to response inhibition across adolescence. *Journal of Abnormal Psychology, 118,* 117–130. http://dx.doi.org/10.1037/a0014657