Personality Disorders: Theory, Research, and Treatment

© 2022 American Psychological Association
ISSN: 1949-2715

2022, Vol. 13, No. 4, 321–324
https://doi.org/10.1037/per0000561

COMMENTARY

Beyond Defending or Abolishing Criterion A: Comment on Morey et al. (2022)

Johannes Zimmermann
Department of Psychology, University of Kassel

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

In their target article, Morey et al. (2022) provided a comprehensive overview of research on Criterion A of the alternative *DSM*–5 model for personality disorders over the past 10 years. Although this overview is undoubtedly informative and helpful, it offers little guidance on both certain weaknesses of Criterion A and research gaps and needs. A structural reason for this could be that the scientific field is divided into 2 camps with respect to Criterion A: Some authors focus mainly on confirmatory results and want to defend Criterion A in its current form, whereas others take the same results as an opportunity to call for its abolition. In this situation, there is little room for a nuanced and constructive discussion of the empirical and conceptual strengths and weaknesses of Criterion A. In this commentary, I use 2 examples to illustrate what research aimed at revising Criterion A might look like. From an empirical point of view, studies that take into account the peculiarities of the latent structure of the Level of Personality Functioning Scale are important because they can reveal local misspecification and challenges for the assessment of single cases. From a conceptual point of view, a more consistent definition of Criterion A in terms of impairments in "capacities" could allow for a more coherent distinction from maladaptive "traits." This commentary hopes to strengthen a discourse that moves beyond the alternatives of defense or abolition toward the concrete improvement of Criterion A.

Keywords: severity, personality functioning, capacity, latent variable models, unfolding item response theory models

The introduction of Criterion A in the alternative *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM–5*), model for personality disorders (AMPD) was undoubtedly a special event in the history of the *DSM* (American Psychiatric Association, 2013). Its inclusion highlighted the importance of the general criteria of personality disorder (PD), elaborated the common denominator of all PDs, defined for the first time what a healthy personality looks like, and made gradual differences in severity directly assessable via the Level of Personality Functioning Scale (LPFS). For the developers of the AMPD, Criterion A represented a way to address the high comorbidity of specific PDs that made the categorical model so inefficient and implausible (Bender et al., 2011). And for psychodynamically oriented researchers, it provided an exciting opportunity to move back into the mainstream of research (Zimmermann et al., 2012).

In their article, Morey et al. (2022) reviewed the history of Criterion A and summarized research findings on the newly developed measures from the past 10 years. In my view, this is a solid and informative review that illustrates that the publication of Criterion A in the AMPD has stimulated a great deal of research and that, overall, the findings can be interpreted in a confirmatory sense. On some issues, the authors perhaps even understate the scope of the knowledge gained so far. For example, we recently summarized results on the interrater reliability of the LPFS (Zimmermann et al., in press). Across 10 studies (N = 676) based on various data sources, the intraclass correlation coefficient of the LPFS total score was .55, 95% confidence interval (CI) [.47, .63]. Crucially, however, nine other studies (N = 276) have been conducted using one of the newly developed structured interviews that were summarized by Morey et al. (2022) in their supplement. Here the intraclass correlation coefficient of the LPFS total score was .83, 95% CI [.75, .92]. This shows that interrater reliability can be considerably improved when using the LPFS based on a tailored structured interview. At the same time, the example of reliability also demonstrates that Morey et al. (2022) failed to point out gaps, uncertainties, and research needs. More specifically, with the exception of

Johannes Zimmermann https://orcid.org/0000-0001-6975-2356
The author would like to thank Elsa Gewehr, Christopher Hopwood, and colleagues at the University of Kassel for helpful feedback on a draft of this article.
Correspondence concerning this article should be addressed to Johannes Zimmermann, Department of Psychology, University of Kassel, Holländische Str. 36-38, Kassel 34127, Germany. Email: jz@uni-kassel.de

two studies with only 33 (Buer Christensen et al., 2018) and 30 participants (Ohse et al., in press), the retest reliability of LPFS ratings has not been investigated. In test–retest designs, the same target person is interviewed and rated independently by two different persons. The corresponding estimate of reliability is usually lower than interrater reliability but more realistic for applied settings (Chmielewski et al., 2015). The existence of only two such studies with small samples indicates that there are still significant gaps in research on the reliability of the LPFS. This exemplifies an overarching pattern in the article by Morey et al. (2022): It is written in the spirit of wanting to confirm, to preserve, to defend. One could argue that this is not surprising because the authors took the lead in developing Criterion A. However, in the current scientific context, it is perhaps appropriate that their role is to tirelessly defend Criterion A against critics, to doggedly explain the ideas behind it, and to collect the evidence in favor of it. In other words, perhaps their attitude could also be understood as a reaction to divisions and camp formations in the scientific field.

This is especially evident in the way "Issues Needing Further Research" is discussed. Instead of pointing out the multiple possible research gaps that researchers should address in the future, this section focuses only on one issue that divides positions in the field —the "Distinction Between Criterion A and Criterion B." It is well-confirmed that in studies in which individuals were examined at a single point in time using a single method, measures of Criterion A and Criterion B are highly correlated, and that in such studies, measures of Criterion A have significant but often low incremental validity compared with measures of Criterion B (Morey et al., 2022; Zimmermann et al., in press). Some researchers are taking these findings as an opportunity to conclude that Criterion A can be dispensed with entirely (Sleep et al., 2019). In contrast, Morey et al. (2022) suggested that one can leave Criterion A as it is and should instead redefine Criterion B in terms of the "normal range" Big Five traits. Both of these positions have merit, but they also show how the scientific field is caught in reflexive positions and counterpositions. Symptomatic of this is that, although the section on "Issues Needing Further Research" covers almost one-third of the article, Morey et al. (2022) gave no specific directions for further research to improve Criterion A. I would like to see a broader discourse in the coming years beyond the alternatives of "Criterion A should remain as is" or "Criterion A should be abolished." What I find lacking in the research on Criterion A, and also in the target article by Morey et al. (2022), is a focus on the empirical inconsistencies on the small scale and on the conceptual fuzziness on the large scale. To effectively coordinate future research and prepare future revisions, a dispassionate recognition of empirical and conceptual weaknesses or gaps seems necessary. In the following, I will discuss two aspects that I find relevant in this regard.

The first point relates to the latent structure of the LPFS. According to *DSM–5*, the 60 prototypical descriptions in Table 2 (American Psychiatric Association, 2013, 775ff) are intended to operationalize a generalized severity continuum. This means that, despite the horizontal differentiation into four different "elements" or "components" of personality functioning (each of which can be further decomposed into three subcomponents according to Table 1) and the vertical differentiation into five different "levels" of impairment, the corresponding empirical ratings would have to satisfy a unidimensional model. Here too there are two conflicting positions that dominate the scientific field and marginalize a perspective interested in corrective revision of Criterion A. On the one hand, critics pointed out that the empirical structure of self-report items developed along the lines of the LPFS, as determined using conventional factor analyses, is multidimensional and not in line with either the components or levels mentioned above (Sleep et al., 2019). On the other hand, Morey et al. (2022) argued that there is a strong general factor and a very high internal consistency of the total score. Again, some consider the LPFS structurally invalid and want to abolish it, whereas others want to leave everything as it is. This pushes more-nuanced and potentially informative patterns of results into the background. For example, when examining the latent structure of the 12 subcomponents of the LPFS, a model with two strongly correlated factors of self and interpersonal functioning usually fits the data quite well, whether in self-reports, informant ratings, or expert ratings based on structured interviews (Zimmermann et al., in press). Although one may argue that this result is consistent with a strong general factor and justifies the formation of an overall score, it suggests that empirically there will be a certain number of individuals who have impairments with respect to either the self or interpersonal relationships and for whom the formation of an overall score may be challenging. A consequence of this finding for a future revision of Criterion A might thus be to establish more-precise rules on how the overall score should be formed or when the threshold for a PD diagnosis is reached (e.g., only when both self *and* interpersonal functioning are at least moderately impaired).

The study of the latent structure should also address the prototypical descriptions within each of the 12 subcomponents. In this regard, conventional factor analyses such as in the article by Sleep et al. (2019) are inappropriate because they cannot adequately model the different levels of impairment. One solution are unfolding item response theory models (Roberts et al., 2000) that allow for examining whether the five descriptions of a given subcomponent are arranged in a theoretically consistent manner along the latent severity continuum. In unfolding item response theory models, a location parameter is estimated for each item, indicating where individuals are located on the latent dimension when they are most likely to match the item description. The only study to date using such models yielded two results that seem relevant here (Zimmermann et al., 2015). First, although most subcomponents were indeed unidimensional, sometimes the pattern of associations turned out to be more complex. For example, we found that the different signs of an impaired "desire and capacity for closeness" that are mentioned in the individual descriptions (e.g., inhibition in Level 1, self-regulatory needs in Level 2, and rejection sensitivity in Level 3) cannot be empirically scaled along a single latent dimension. Second, the theoretically hypothesized severity levels and the empirically estimated location parameters were highly correlated across items, largely supporting the severity assignments of the LPFS descriptions. Some location parameters however also emerged that deviated from this general pattern. For example, the location parameters of the moderate, severe, and extreme impairment items in some subcomponents (e.g., "self-esteem" and "self-reflective functioning") were all uniformly at the dysfunctional pole of the latent continuum, suggesting that, on average, raters did not capture the subtle differences in severity that the descriptions were intended to convey. If one is interested in optimizing the LPFS in an evidence-based manner, further such studies are essential.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

The second point takes me back to the "Distinction Between Criterion A and Criterion B." Here too it might be fruitful to look beyond the alternatives of abolishing or defending Criterion A toward possibilities for conceptual clarification. In my view, it would be helpful to develop a better rationale for how and why the phenomena currently described in Criteria A and B should be distinguished from one another (Zimmermann et al., 2015). Various opinions have been expressed about the conceptual differences between Criterion A and Criterion B, including their focus on severity versus style (Morey et al., 2022), their emphasis on processes versus structure (Pincus et al., 2020), and their origins in psychodynamic versus trait psychological traditions and methods. I would like to raise here another possibility: One could finally take seriously the rather cursory designation of phenomena subsumed under Criterion A as "capacities" (Morey et al., 2022), thereby unambiguously distinguishing them from maladaptive traits in Criterion B. Capacities (e.g., empathy) refer to how well someone can exhibit certain behaviors when they are motivated and the situation calls for it. Maladaptive traits (e.g., callousness) refer to how typical it is for a person not to exhibit the corresponding behaviors or to exhibit them in an inappropriate manner. As with corresponding skills and traits (Soto et al., 2021), corresponding impairments in capacities and maladaptive traits are probably positively correlated: Someone who is not good at something will do it less often or fail at it more often; someone who does something infrequently or approaches it incorrectly will not get better at it. Thus, with such a conceptualization, some degree of substantive overlap and empirical correlation between Criterion A and Criterion B might be warranted.

Interestingly, this would also suggest a causal relationship between the two criteria: Maladaptive traits are more likely the expression or consequence of impaired abilities than vice versa (e.g., a person tends to be callous because their capacity for empathy is impaired; Zimmermann et al., 2015). Such an understanding would be in line with the view of Sharp and Wall (2021), who pointed out that the LPFS does not map the "consequences" of personality pathology but rather their "source." In this sense, Criterion A would indeed represent the "return of theory" (Pincus, 2011) to a formerly atheoretical classification system: It would be tantamount to adopting an ability-deficit hypothesis for PD that assumes that all PDs are necessarily grounded in structural impairments (cf. OPD Task Force, 2008). This emphasis would also be in line with the harmful dysfunction concept for mental disorders posited by Wakefield (1992): As long as PD has no reference to an internal dysfunction, it is not actually a "disorder" but merely a pattern of experience and behavior that is harmful in terms of social values. The reference to impaired capacities would provide the reference to an underlying dysfunction in which a psychological mechanism fails and no longer performs the natural function for which it was selected in the course of evolution.

It would be stimulating to reflect on the consequences of this conceptual sharpening for diagnostic and therapeutic practice. For example, in this case, it would seem advisable to develop a test battery with maximum performance tests that capture emotional, self-related, and interpersonal competencies (Jauk & Ehrenthal, 2021; Leising et al., 2011; Olderbak & Wilhelm, 2020). And one could examine currently available treatment approaches for PDs to see how they already stabilize, make available, and train various basic capacities as defined in Criterion A. It should however also be noted that not all phenomena currently included in Criteria A and B conform to the classification of capacities versus traits. For example, in Criterion A, not only is a "capacity for closeness" defined but also a "desire for closeness"—suggesting that a lack or deadlock of universal motives may also play a role in some PDs. And in the definitions of Criterion B, explicit references to ability deficits are occasionally present (e.g., "anhedonia" is defined as "deficits in the capacity to feel pleasure"; American Psychiatric Association, 2013, p. 779). Such conceptual sharpening would thus also involve adjustments to the content of the AMPD and could introduce unforeseen new problems. My point here is not to make a case for this solution. In fact, I was already involved in a very different proposal to separate the description of personality traits from the evaluation of their negative consequences for the purpose of diagnosing PD (Leising & Zimmermann, 2011). My main point here is to show that, beyond the alternative between abolishing or defending Criterion A, there are ideas about how to revise the AMPD in a constructive way and thereby stimulate new research.

For those interested in what has been accomplished so far in research on Criterion A, the article by Morey et al. (2022) provides a very useful overview. It serves its purpose as a celebratory publication for the anniversary of Criterion A, and it highlights the authors' great achievement in introducing Criterion A into *DSM–5*, thereby inspiring many researchers in the field. However, for those who want to know what has not yet been achieved in research on Criterion A, what open questions exist, and what aspects of Criterion A need improvement, the article falls short. I have singled out two aspects that I believe deserve more attention if Criterion A is to be made more empirically viable and conceptually coherent. But the wish list for future research on AMPD is of course longer. It will be critical in the future to move beyond the currently dominant study designs in which individuals are studied at one point in time and with one method and from which, unfortunately, the variance of the construct, method, and situation cannot be separated (Zimmermann et al., 2019). And if one really wants to convincingly demonstrate the clinical utility of the AMPD, further case examples or consumer surveys will not help. Then it will be a matter of using randomized controlled trials to show that using the AMPD versus the categorical model of PD in clinical practice really makes a tangible difference for patients and practitioners (Lewis et al., 2019).

References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5* (5th ed.). Author.

Bender, D. S., Morey, L. C., & Skodol, A. E. (2011). Toward a model for assessing level of personality functioning in *DSM-5*, part I: A review of theory and methods. *Journal of Personality Assessment*, *93*(4), 332–346. https://doi.org/10.1080/00223891.2011.583808

Buer Christensen, T., Paap, M. C. S., Arnesen, M., Koritzinsky, K., Nysaeter, T.-E., Eikenaes, I., Germans Selvik, S., Walther, K., Torgersen, S., Bender, D. S., Skodol, A. E., Kvarstein, E., Pedersen, G., & Hummelen, B. (2018). Interrater reliability of the structured clinical interview for the *DSM-5* alternative model of personality disorders module I: Level of personality functioning scale, *Journal of Personality Assessment*, *100*(6), 630–641. https://doi.org/10.1080/00223891.2018.1483377

Chmielewski, M., Clark, L. A., Bagby, R. M., & Watson, D. (2015). Method matters: Understanding diagnostic reliability in *DSM–IV* and *DSM–5*. *Journal of Abnormal Psychology*, *124*(3), 764–769. https://doi.org/10.1037/abn0000069

Jauk, E., & Ehrenthal, J. C. (2021). Self-reported levels of personality functioning from the operationalized psychodynamic diagnosis (OPD) system and emotional intelligence likely assess the same latent construct. *Journal of Personality Assessment*, *103*(3), 365–379. https://doi.org/10.1080/00223891.2020.1775089

Leising, D., Krause, S., Köhler, D., Hinsen, K., & Clifton, A. (2011). Assessing interpersonal functioning: Views from within and without. *Journal of Research in Personality*, *45*(6), 631–641. https://doi.org/10.1016/j.jrp.2011.08.011

Leising, D., & Zimmermann, J. (2011). An integrative conceptual framework for assessing personality and personality pathology. *Review of General Psychology*, *15*(4), 317–330. https://doi.org/10.1037/a0025070

Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., Hoffman, M., Scott, K., Lyon, A., Douglas, S., Simon, G., & Kroenke, K. (2019). Implementing measurement-based care in behavioral health: A review. *JAMA Psychiatry*, *76*(3), 324–335. https://doi.org/10.1001/jamapsychiatry.2018.3329

Morey, L. C., McCredie, M. N., Bender, D. S., & Skodol, A. E. (2022). Criterion A: Level of personality functioning in the alternative *DSM-5* model for personality disorders. *Personality Disorders: Theory, Research, and Treatment*, *13*(4), 305–315. https://doi.org/10.1037/per0000551

Ohse, L., Zimmermann, J., Kerber, A., Kampe, L., Mohr, J., Kendlbacher, J., Busch, O., Rentrop, M., & Hörz-Sagstetter, S. (in press). Reliability, structure and validity of Module I (personality functioning) of the structured clinical interview for the alternative *DSM–5* model for personality disorders (SCID-5-AMPD-I). *Personality Disorders: Theory, Research, and Treatment*.

Olderbak, S., & Wilhelm, O. (2020). Overarching principles for the organization of socioemotional constructs. *Current Directions in Psychological Science*, *29*(1), 63–70. https://doi.org/10.1177/0963721419884317

OPD Task Force. (2008). *Operationalized psychodynamic diagnosis OPD–2: Manual of diagnosis and treatment planning*. Hogrefe & Huber.

Pincus, A. L. (2011). Some comments on nomology, diagnostic process, and narcissistic personality disorder in the *DSM-5* proposal for personality and personality disorders. *Personality Disorders: Theory, Research, and Treatment*, *2*(1), 41–53. https://doi.org/10.1037/a0021191

Pincus, A. L., Cain, N. M., & Halberstadt, A. L. (2020). Importance of self and other in defining personality pathology. *Psychopathology*, *53*(3-4), 133–140. https://doi.org/10.1159/000506313

Roberts, J. S., Donoghue, J. R., & Laughlin, J. E. (2000). A general item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement*, *24*(1), 3–32. https://doi.org/10.1177/01466216000241001

Sharp, C., & Wall, K. (2021). *DSM-5* level of personality functioning: Refocusing personality disorder on what it means to be human. *Annual Review of Clinical Psychology*, *17*(1), 313–337. https://doi.org/10.1146/annurev-clinpsy-081219-105402

Sleep, C. E., Lynam, D. R., Widiger, T. A., Crowe, M. L., & Miller, J. D. (2019). An evaluation of *DSM-5* Section III personality disorder Criterion A (impairment) in accounting for psychopathology. *Psychological Assessment*, *31*(10), 1181–1191. https://doi.org/10.1037/pas0000620

Soto, C. J., Napolitano, C. M., & Roberts, B. W. (2021). Taking skills seriously: Toward an integrative model and agenda for social, emotional, and behavioral skills. *Current Directions in Psychological Science*, *30*(1), 26–33. https://doi.org/10.1177/0963721420978613

Wakefield, J. C. (1992). The concept of mental disorder. On the boundary between biological facts and social values. *American Psychologist*, *47*(3), 373–388. https://doi.org/10.1037/0003-066X.47.3.373

Zimmermann, J., Böhnke, J. R., Eschstruth, R., Mathews, A., Wenzel, K., & Leising, D. (2015). The latent structure of personality functioning: Investigating Criterion A from the alternative model for personality disorders in *DSM-5*. *Journal of Abnormal Psychology*, *124*(3), 532–548. https://doi.org/10.1037/abn0000059

Zimmermann, J., Ehrenthal, J. C., Cierpka, M., Schauenburg, H., Doering, S., & Benecke, C. (2012). Assessing the level of structural integration using operationalized psychodynamic diagnosis (OPD): Implications for *DSM-5*. *Journal of Personality Assessment*, *94*(5), 522–532. https://doi.org/10.1080/00223891.2012.700664

Zimmermann, J., Hopwood, C. J., & Krueger, R. F. (in press). The *DSM-5* level of personality functioning scale. In R. F. Krueger & P. H. Blaney (Eds.), *Oxford textbook of psychopathology* (4th ed.). Oxford University Press.

Zimmermann, J., Kerber, A., Rek, K., Hopwood, C. J., & Krueger, R. F. (2019). A brief but comprehensive review of research on the alternative *DSM-5* model for personality disorders. *Current Psychiatry Reports*, *21*(9), Article 92. https://doi.org/10.1007/s11920-019-1079-z

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.