

The Construct Behind Criterion A

Why the field keeps confusing a developmental construct with its measurement tools



KIMBERLY (KIMMIE) GILBERT

MAR 08, 2026

If you spend any time in the personality disorders literature, you'll eventually notice something strange: people use the phrase "personality functioning" to mean at least three completely different things, often in the same paragraph, without seeming to notice.

There's personality functioning as a *psychological construct*: the degree to which a person has developed, through the interaction of temperament and early relational experience, the capacity to maintain a coherent sense of self and to engage in meaningful relationships with others. This idea has deep roots in psychodynamic and developmental theory stretching back decades.

There's the *Level of Personality Functioning* (LPF) as written in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5), Section III, a specific four-domain operationalization (Identity, Self-Direction, Empathy, Intimacy) rated on a severity continuum from 0 to 4.

And then there are *self-report questionnaires* designed to capture the LPF through item endorsement, instruments like the LPFS-Self Report (LPFS-SR) or the LPFS-Brief Form (LPFS-BF).

These are not the same thing, and I don't think the distinctions between them get enough airtime.

A lot of confusion in the literature (and in conversations I've had with colleagues) traces back to people talking past each other because

they're using the same phrase to refer to different levels of the concept. I think it's worth laying out what each one is and why the differences matter.

The Construct Has a History Worth Knowing

One of the frustrations I have with how personality functioning (PF) gets discussed is that the conversation often starts and ends with the Alternative Model of Personality Disorders (AMPD). In reality, the ideas behind PF have a much longer intellectual history. The AMPD didn't invent the concept; it operationalized a set of developmental capacities that multiple traditions in psychology had already been describing for decades.

Across different theoretical schools, researchers repeatedly converged on a similar insight: psychological health depends on the development of core capacities involving the self and relationships with others.

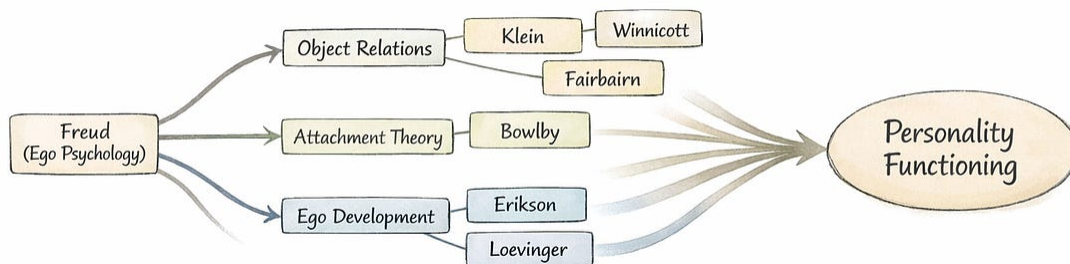
Several traditions contributed to this idea.

First, psychoanalytic and object relations models described how early relationships shape the structure of the self and the way people experience others. Freud's early structural thinking about the ego laid the groundwork for later models of psychological capacities. Object relations theorists such as Klein (1946), Fairbairn (1949), and Winnicott (1958, 1965) focused on how internal representations of self and others develop through early caregiving experiences. Kohut's (1971) self psychology emphasized how mirroring and idealization support the development of stable self-esteem. Later, Kernberg's model of personality organization (1967, 2004) articulated how identity integration, defensive operations, and reality testing differentiate levels of personality pathology.

Second, developmental models of identity and ego growth emphasized how the self becomes more integrated across the lifespan. Erikson (1963) framed identity formation as a central developmental task extending from adolescence into adulthood, while Loevinger (1976) described ego development as a progression toward increasingly complex and integrated ways of understanding oneself and others.

Third, attachment and mentalization frameworks highlighted the relational origins of these capacities. Bowlby's (1969) attachment theory proposed that early caregiving relationships create internal working models that guide expectations about self and others. Mentalization theory later elaborated how the capacity to understand behavior in terms of mental states develops within attachment relationships (Fonagy et al., 2002).

Blatt's (1974, 2008) work brought these strands together by identifying two intertwined lines of psychological development: relatedness and self-definition. These two dimensions map closely onto the distinction between interpersonal and self-functioning that later became central to the Level of Personality Functioning Scale.



In this sense, PF didn't emerge from a single theory. It sits at the convergence of traditions that all arrived at a similar conclusion: there is a set of core developmental capacities that determine how well someone can function as a person in the world. That's a fundamentally

different intellectual project than tallying trait scores or counting symptom criteria.

The Incremental Validity Question

One of the most common research designs in the AMPD literature involves testing whether personality functioning measures add predictive value beyond trait measures (or vice versa). You take a self-report PF measure, throw it into a regression model alongside the Big Five or PID-5 traits, and report whether it accounts for additional variance in some outcome. That work mattered and I don't want to dismiss it. In fact, I have a paper on the AMPD that used incremental validity analyses to account for variance in borderline features (Gilbert et al., in press).

The deeper issue is that the incremental validity framework implicitly treats PF as something that needs to *justify its existence* relative to trait models. But from the perspective of the AMPD itself, PF and traits are designed to do different jobs (Morey et al., 2022; Hopwood, 2025). PF captures the overall severity of personality pathology. Traits capture the style in which that pathology tends to manifest. Asking which one predicts outcomes better misunderstands the architecture of the model. PF isn't a general factor extracted from the covariance among maladaptive indicators. It's a theoretical claim about developmental capacity: how far along someone has gotten in the process of building a stable self and the ability to relate meaningfully to others. Traits describe how that level of development expresses itself in characteristic patterns of thought, feeling, and behavior. When you pit them against each other in a regression, of course there's massive overlap. The aggregate of all the specific ways someone is struggling will naturally approximate their overall level of developmental impairment. That's not evidence of redundancy. That's the model working as intended.

The Measurement Challenge

Now, one of the reasons these distinctions matter practically is that measuring PF well is genuinely hard. The construct is inherently psychodynamic. It involves processes that are partly unconscious, relational, and context-dependent. That makes it difficult to capture through self-report. You're asking people to reflect on capacities they may, by definition, lack the self-awareness to evaluate. There's something almost paradoxical about handing someone a questionnaire and asking, "How good are you at understanding your own mind?" The people who are worst at it are the least likely to know.



To make this concrete: imagine a patient sitting in your office who, when asked about a recent conflict with his partner, can only describe

what happened in terms of what the other person did wrong. He can't access his own role in the dynamic, can't hold a stable image of his partner as someone who is both frustrating and loving, and oscillates between idealizing and devaluing her depending on the last interaction. In a clinical interview, a skilled assessor picks up on this immediately. It shows up in how the patient talks, not just what he endorses. On a questionnaire, this same patient might rate himself as having decent self-awareness, because from inside the experience, he genuinely believes he does. The construct is visible in the relational process in ways that item endorsement struggles to capture.

There are interview-based approaches that do better here (e.g. the Structured Interview of Personality Organization-Revised [STIPO-R], the Semi-Structured Interview for Personality Functioning DSM-5 [STiP-5.1], the OPD clinical interview), but they don't scale easily. And the self-report measures that do scale tend to have items that look a lot like trait items and problem-in-living items (McCabe et al., 2021). That's a big part of why PF and traits are so hard to pull apart in cross-sectional questionnaire data.

That's not a failure of the researchers doing this work. It's a genuine constraint of the method. Nearly every questionnaire item, no matter what construct it's designed to measure, will reflect some mix of latent disposition and manifest adaptation.

It's also worth noting that the broader PF construct, as understood through its psychodynamic predecessors, includes domains that the LPFS doesn't foreground. Things like the maturity of defensive operations, the quality of reality testing under stress, and the role of aggression in personality organization are all central to frameworks like Kernberg's model and the OPD system. There's evidence from the assessment literature that some of these domains predict clinically important outcomes (like self-harm and suicidality) even after controlling for overall severity (Kampe et al., 2018; Lowyck et al., 2013). None of this means the LPFS got it wrong; it was designed to be

broadly accessible and to capture the core of the construct without requiring deep immersion in psychodynamic theory, and it succeeds at that. But it does mean the LPFS is one operationalization of a broader construct, and researchers and clinicians working with PF should be aware of the dimensions it foregrounds and the ones it leaves in the background.

But the difficulty of measuring something well doesn't mean the thing itself is flawed. It means our methods haven't caught up to the phenomenon. And I think that's a project worth investing in: finding rigorous ways to make psychodynamic constructs measurable without gutting their explanatory power. That aligns with what I see as the field's most important unfinished business, moving beyond purely descriptive models of personality pathology toward genuinely mechanistic and explanatory thinking, especially when it comes to things like identity, self-concept, and the internalization of relational experience.

Not Anti-Trait, But Beyond Trait

This measurement difficulty is itself part of why PF keeps getting collapsed into trait space. When the instruments used to measure PF produce items that behave like trait items, it's easy to conclude that the constructs are redundant. But that conclusion follows from the properties of the measurement, not the properties of the phenomenon. The question worth asking isn't "do these questionnaires produce distinguishable scores?" but "are these constructs doing different explanatory work?" I think they clearly are.

I want to be clear, because this gets polarized fast: none of this is an anti-trait position. Traits capture real and important variance in personality pathology.

But traits tell you *what* someone tends to do. PF asks *why*, and *with what* psychological resources, they do it. Consider two patients who both

score high on negative affectivity and low on agreeableness. One has a reasonably coherent sense of identity and some capacity for self-reflection, even if she struggles with emotional reactivity. The other is operating from a fragmented self-structure, using primitive defenses (e.g. splitting or projection), unable to hold a stable image of herself or others from one session to the next. The trait profile is identical. The clinical reality is completely different. That's the gap PF is supposed to fill.

Or think about it from the diagnostic side. It's common in clinical settings to encounter someone who doesn't check enough boxes for any single categorical personality disorder but who is clearly struggling with the basics. Unstable self-esteem. Shallow and exploitative relationships. Difficulty tolerating ambivalence. Reliance on splitting and projection to manage distress. Under the old system, this person technically doesn't have a personality disorder. Under a dimensional framework anchored in PF, the impairment is obvious and the treatment implications are immediate. You don't need five of nine criteria to know that someone's capacity to hold a coherent image of themselves and others is compromised. That's something a clinician can see, name, and work with, and it cuts across whatever categorical labels happen to fit or not fit on a given day.

The p-Factor Connection

There's another reason to take PF seriously as a construct: it may offer meaningful content to the much-debated general factor of psychopathology.

The p-factor has been a hot topic for over a decade, the idea that there's a single underlying dimension of severity running through all forms of mental illness. But it's also attracted serious criticism. Some researchers have argued that it may be more statistical artifact than

substantive entity, that its content shifts from study to study, loading heavily on psychosis in one sample, on depression and anxiety in another (Watts et al., 2023). Without a clear theoretical account of what the general factor actually *is*, it risks being an empty summary statistic.

PF could provide that theoretical content. If what the p -factor really captures is variation in core psychological capacities, the ability to maintain a stable sense of self, to regulate emotion, to understand others as separate minds, to direct one's own behavior toward meaningful goals, then PF gives us a framework for understanding *why* psychopathology severity cuts across diagnostic categories. Some recent large-scale work has suggested that PF does in fact account for a large share of the variance in higher-order psychopathology dimensions (Kerber et al., 2024; see also Wright et al., 2016). That's suggestive, and it points toward a version of the general factor that actually means something.

Where Do We Go?

The open questions are clear enough. We need longitudinal and process-level research on how PF interacts dynamically with traits, relational patterns, and symptom trajectories, not just cross-sectional regressions. We need studies that examine PF as a moderator and mediator of treatment across theoretical orientations, including digital and blended interventions. We need integration of PF into clinical training so that therapists of any stripe can use it to understand what's happening when a patient can't tolerate ambivalence, keeps enacting destructive relational cycles, or experiences the therapist as a threat rather than a resource.

We also need to get serious about multimethod assessment. If the argument of this post is right, that PF is a developmental construct poorly captured by self-report, then the field can't keep relying almost

exclusively on questionnaires to study it. The psychoanalytic tradition has always emphasized that personality functioning reveals itself in how people talk about themselves and others, not just what they endorse on a checklist. An instrument like the STIPO-R captures the quality of someone's internal representations and the maturity of their defenses through extended clinical interview. The Reflective Functioning Scale codes mentalization quality from Adult Attachment Interview (AAI) transcripts, assessing reflective functioning—the capacity to interpret behavior in terms of underlying mental states by rating the frequency, complexity, and coherence of reflective statements about early relationships. These approaches get at the construct in ways no Likert scale can. But they require trained clinicians, hours of interview time, and expert coding, which is why they rarely show up in large-scale research.

This is where narrative and text-based methods become especially interesting.

For decades, psychodynamic researchers have relied on narrative assessment traditions to study constructs that don't show up well in self-report. Rather than asking people to endorse statements on a Likert scale, these approaches analyze how individuals talk about themselves and others in extended interviews. The underlying assumption is simple: psychological capacities reveal themselves in the structure and content of a person's narrative. Coherence, complexity, emotional nuance, and the ability to hold ambivalent representations of self and others all become visible in how someone tells their story.

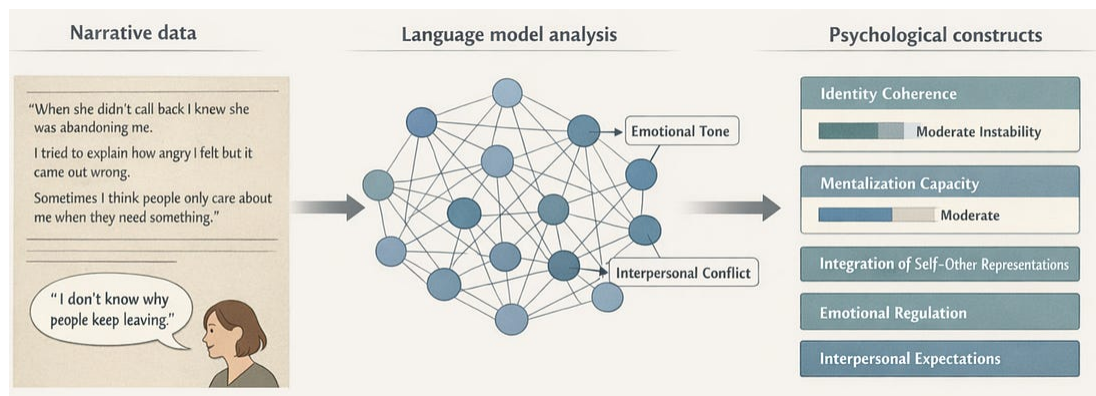
To study these features systematically, researchers developed **coding systems** applied to interview transcripts and clinical material. The Reflective Functioning Scale, for example, assesses mentalization by coding Adult Attachment Interview transcripts for the presence, frequency, and sophistication of reflective statements about mental states. Other systems analyze object relations themes, narrative coherence, or the complexity of representations of self and others.

These methods get much closer to the phenomenon than questionnaire items can.

The problem has always been *scalability*. Coding narrative data requires trained raters, extensive reliability training, and hours of careful evaluation for each interview transcript. As a result, these methods rarely appear in large-scale studies, even though they arguably capture the construct more directly than self-report instruments.

Recent advances in large language models (LLMs) may offer a way around this bottleneck. These systems are increasingly capable of identifying many of the same narrative features human coders are trained to detect: coherence, affective tone, representational complexity, the presence of mental state reasoning, and the ability to hold contradictory perspectives about self and others. If an LLM can reliably approximate what trained raters do when coding reflective functioning from an AAI transcript, or what clinicians assess across domains in a STIPO-style interview, then large-scale analysis of narrative data becomes possible.

In effect, this would open the door to something the field has struggled to achieve for decades: computational approaches to psychodynamic assessment. Instead of collapsing developmental constructs like personality functioning into self-report items that resemble trait measures, researchers could analyze how people actually talk about themselves and their relationships. This isn't a solved problem yet. But it's a tractable one, and it could fundamentally expand the kinds of data the field can bring to bear on questions about personality development and pathology.



And we need to get clearer about the relationship between the construct and its operationalizations. Personality functioning is bigger than the LPFS. It's bigger than any one questionnaire or interview protocol. **It's a way of thinking about what makes people psychologically capable or fragile, rooted in decades of developmental and psychodynamic theory.** The LPFS succeeded in bringing that thinking into the mainstream diagnostic conversation. The next challenge is making sure that, in the process of operationalizing it, we don't flatten the very phenomenon it was meant to capture.

References

- 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.
- Blatt, S. J. (1974). Levels of object representation in anaclitic and introjective depression. *Psychoanalytic Study of the Child*, 29, 107–157.
- Blatt, S. J. (2008). *Polarities of experience: Relatedness and self-definition in personality development, psychopathology, and the therapeutic process*. American Psychological Association.
- Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. Basic Books.

Erikson, E. H. (1963). *Childhood and society* (2nd ed.). Norton.

Fairbairn, W. R. D. (1949). Steps in the development of an object-relations theory of the personality. *British Journal of Medical Psychology*, 22(1–2), 26–31.

Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self*. Other Press.

Freud, S. (1905). Three essays on the theory of sexuality. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 7, pp. 125–243). Hogarth Press.

Hopwood, C. J. (2025). Personality functioning, problems in living, and personality traits. *Journal of Personality Assessment*, 107(2), 143–158.

Hörz-Sagstetter, S., Ohse, L., & Kampe, L. (2021). Three dimensional approaches to personality disorders: A review on personality functioning, personality structure, and personality organization. *Current Psychiatry Reports*, 23, 45.

Kampe, L., Zimmermann, J., Bender, D., Caligor, E., Borowski, A.-L., Ehrental, J. C., Benecke, C., & Hörz-Sagstetter, S. (2018). Comparison of the Structured DSM-5 Clinical Interview for the Level of Personality Functioning Scale with the Structured Interview of Personality Organization. *Journal of Personality Assessment*, 100(6), 642–649.

Kerber, A., Ehrental, J. C., Zimmermann, J., Remmers, C., Nolte, T., Wendt, L. P., Heim, P., Müller, S., Beintner, I., & Knaevelsrud, C. (2024). Examining the role of personality functioning in a hierarchical taxonomy of psychopathology using two years of ambulatory assessed data. *Translational Psychiatry*, 14(1), 340.

Kernberg, O. F. (1967). Borderline personality organization. *Journal of the American Psychoanalytic Association*, 15(3), 641–685.

Kernberg, O. F. (2004). Borderline personality disorder and borderline personality organization: Psychopathology and psychotherapy. In J. J. Magnavita (Ed.), *Handbook of personality disorders: Theory and practice* (pp. 92–119). John Wiley & Sons.

Klein, M. (1946). *Envy and gratitude: The writings of Melanie Klein*. Hogarth Press and The Institute of Psychoanalysis.

Kohut, H. (1971). *The analysis of the self: A systematic approach to the psychoanalytic treatment of narcissistic personality disorders*. University of Chicago Press.

Loevinger, J. (1976). *Ego development: Conceptions and theories*. Jossey-Bass.

Lowyck, B., Luyten, P., Verhaest, Y., Vandeneede, B., & Vermote, R. (2013). Levels of personality functioning and their association with clinical features and interpersonal functioning in patients with personality disorders. *Journal of Personality Disorders*, 27(3), 320–336.

McCabe, G. A., Oltmanns, J. R., & Widiger, T. A. (2021). Criterion A scales: Convergent, discriminant, and structural relationships. *Assessment*, 28(3), 813–828.

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.

Watts, A. L., Greene, A. L., Bonifay, W., & Fried, E. I. (2023). A critical evaluation of the p-factor literature. *Nature Reviews Psychology*.

Winnicott, D. W. (1958). The capacity to be alone. *International Journal of Psycho-Analysis*, 39, 416–420.

Winnicott, D. W. (1965). *The maturational processes and the facilitating environment: Studies in the theory of emotional development*. International Universities Press.

Wright, A. G. C., Hopwood, C. J., Skodol, A. E., & Morey, L. C. (2016). Longitudinal validation of general and specific structural features of personality pathology. *Journal of Abnormal Psychology, 125*(8), 1120–1134.

Author note: This post was inspired by a recent conversation I had with Aidan Wright. He laid out very clearly the distinction between PF as a construct, the DSM-5 LPF, and LPF-based measurements better than I could articulate so I thank him.
