Clarifying Measurement and Construct-Level Inference in Myers and Abramowitz's Review of the Inference-Based Approach

Frederick Aardema [1,2,3]

[1] Montreal Mental Health University Institute Research Center, Montreal, Canada
[2] Department of Psychiatry and Addictology, University of Montreal, Montreal, Canada
[3] Obsessive Compulsive Disorders Clinical Study Center, Montreal, Canada

Introduction

Myers and Abramowitz (2025) provide a critical review of the inference-based approach (IBA) documenting consistent associations between inferential confusion (IC) with symptoms of obsessive–compulsive disorder (OCD), alongside evidence of specificity, strong internal consistency, and a unidimensional structure.

At the same time, the review raises measurement-level concerns that qualify the evidentiary base for IC, but several interpretive moves conflate limitations of specific operationalizations with limitations of the construct itself. A more conservative interpretation recognizes measurement constraints without undermining IC as a reasoning process.

Construct breadth, phenomenology, and self-report

A central concern raised in the review is that IC may be "broad" or "nebulous," and that existing instruments may not capture its "entirety." Many established OCD constructs are similarly multifaceted and require multiple operationalizations without invalidating the construct, a point long recognized in the construct-validation literature (Cronbach & Meehl, 1955; Campbell & Fiske, 1959).

Relatedly, the review suggests that IC measures rely too heavily on phenomenological self-report and may therefore reflect OCD-specific experience rather than inferential processes per se. This objection rests on a misunderstanding of how IC is defined within IBA. IC is conceptualized as a disturbance in lived reasoning, making phenomenological expression constitutive rather than contaminating.

The review further questions whether inferential processes are accessible to introspection. Such concerns align with known limitations of insight in OCD (Neziroglu et al., 1999), but reduced insight would be expected to attenuate, rather than inflate, self-reported differentiation. The consistent observation of systematic associations between IC measures and OC symptom severity indicates that self-report captures meaningful variance, even if imperfectly. Recent work further shows that insight itself can be assessed via self-report (Ouellet-Courtois, Audet, & Aardema, 2024).

Scope of critique and evidentiary coverage

Although framed as a critical review of IBA, several of the review's central conclusions regarding construct validity are derived primarily from concerns about the wording, response formats, and endorsement properties of IC questionnaires. This focus is informative, but not comprehensive. Several components central to the approach, particularly treatment outcome studies and work integrating IC with fear-of-self constructs, receive limited attention. Importantly, the review excludes intervention studies by scope, treating outcome and process evidence as conceptually irrelevant to IC.

Relatedly, the review concludes that causal evidence for IBA is confined to a single experimental study. While direct laboratory manipulation of IC remains limited, this conclusion reflects a narrow evidentiary frame. Although outcome studies do not substitute for experimental manipulation, longitudinal designs contribute to construct by by situating IC within a broader nomological network that includes temporal coherence and clinically responsive change.

In this context, excluding outcome and process studies applies a more restrictive evidentiary standard than is typically required for construct validation. When considered alongside the omission of fear-of-self research, the resulting evidentiary base is insufficiently broad to sustain the construct-level conclusions advanced.

Endorsement-based item screening and construct inference

The review places particular weight on endorsement-based item screening in the development of a single questionnaire, the Inferential Confusion Questionnaire–Expanded Version (ICQ-EV; Aardema et al., 2010), suggesting that removal of low-endorsement items may have reshaped the measure toward OCD phenomenology and rendered its associations tautological. This inference warrants closer scrutiny.

Endorsement-based screening is a standard psychometric procedure for addressing floor effects and restricted variance (Nunnally & Bernstein, 1994; DeVellis, 2017). Items were removed on the basis of frequency rather than symptom correlation. Claims of tautology or construct distortion depend on item-level evidence indicating that removed items uniquely indexed inferential mechanisms or that retained items restated symptom content. Absent such evidence, the inference regarding construct validity remains speculative and selective.

Consistent with this, earlier versions of the ICQ, developed using different item-generation and selection strategies, show similar factor structures and associations with OC symptoms. If endorsement-based pruning were driving spurious coherence, divergence across versions would be expected. That pattern is not evident.

Generality, specificity, and convergence across measures

The review argues that generalized ICQ phrasing conflicts with selective deployment, while task-based measures rely on fixed scenarios that may not fit idiosyncratic obsessions. These objections pull in opposite directions. Generality trades off with contextual specificity, and realism with universality. This trade-off is common in psychopathology research and does not imply conceptual inconsistency within IBA.

Against this background, the review characterizes moderate to strong correlations between questionnaire-based and task-based measures of IC as only "modest" evidence of convergent validity. From a construct validation perspective, however, convergence between independently developed measures that differ in structure, format, and response demands constitutes meaningful evidence that they are tapping a shared underlying process.

The review further acknowledges that measures of IC often demonstrate incremental validity alongside obsessive belief measures in predicting OC symptoms, but questions the effect size. Clearer reporting of semi-partial correlations and incremental R^2 is warranted. However, incremental variance is a limited test of theoretical validity when constructs are hierarchically related (Meehl, 1990).

Conclusion

Myers and Abramowitz (2025) provide a methodologically detailed review that emphasizes measurement rigor, replication, and evidentiary standards. Importantly, they also characterize IC as a "genuine signal" whose effects may be obscured by methodological noise. However, several of the review's subsequent inferences extend beyond what the available data warrant. Measurement-level limitations are treated as substantial qualifications of the construct itself; endorsement-based item screening is framed as a validity threat despite psychometric precedent; phenomenology is construed as contaminating rather than constitutive; and scope-based exclusions are treated as neutral rather than inferentially constraining.

A more conservative interpretation is that IC remains a coherent reasoning process implicated in obsessive doubt, supported by convergent evidence across measures and methods. As with other cognitive or behavioral constructs in OCD research, continued conceptual refinement and improved operational precision are both expected and welcome, and do not, in themselves, undermine construct validity. While further research is needed to clarify its scope and causal role, an issue Myers and Abramowitz note is not unique to IBA, construct validity is not typically discounted on this basis. Continued progress will therefore depend on multi-method assessment and direct tests of inferential processes, without conflating the limits of current operationalizations with the viability of the construct itself.

References

Aardema, F., Wu, K. D., Careau, Y., O'Connor, K., Julien, D., & Dennie, S. (2010). The expanded version of the Inferential Confusion Questionnaire: Further development and validation in clinical and non-clinical samples. *Journal of Psychopathology and Behavioral Assessment, 32*, 448-462.

Baraby, L-P., Bourguignon, L., & Aardema, F. (2022). The relevance of dysfunctional reasoning to OCD and its treatment: Further evidence for inferential confusion utilizing a new task-based measure. *Journal of Behavior Therapy and Experimental Psychiatry*, 101728.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56*, 81–105.

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Thousand Oaks, CA: Sage Publications.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281–302.

Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports*, *66*, 195–244.

Myers, N.S. & Abramowitz, J.S. (2025). Unpacking inferential confusion: A critical review of the inference-based approach to obsessive-compulsive disorder. *Journal of Obsessive-Compulsive and Related Disorders, 47*, 2025, 100983.

Neziroglu, F., McKay, D., Yaryura-Tobias, J. A., Stevens, K. P., & Todaro, J. (1999). The Overvalued Ideas Scale: Development, reliability, and validity in obsessive–compulsive disorder. *Behaviour Research and Therapy, 37*(9), 881–902.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Ouellet-Courtois, C., Audet, J.-S., & Aardema, F. (2024). The Cognitive Obsessional Insight Scale (COGINS): A new measure of cognitive insight in obsessive–compulsive and related disorders. *Journal of Cognitive Psychotherapy: An International* Quarterly, 38, 133-156.