A NEW PERSPECTIVE ON MALADAPTIVE PERSEVERATION: THE IMPACT OF UNFALSIFIABLE BELIEF SYSTEMS ON INSIGHTFUL REASONING

Stijn Van Severen*
Department of Experimental Psychology, Ghent University, Ghent, Belgium

Marc Brysbaert
Department of Experimental Psychology, Ghent University, Ghent, Belgium

Abstract

Belief systems that resist contradictory evidence can make individuals less open to changing their perspective, possibly affecting their ability to effectively navigate complex problem domains. To explore this, 60 participants took part in a 55-minute experiment where they temporarily adopted either refutable or irrefutable reasoning styles prior to insight tasks. After exposure to rigid cognitive frameworks, participants had more difficulty with insightful reasoning where representational change was a prerequisite for finding the solution, as indicated by an overall increase in reaction time. A second analysis found that participants who were more drawn to irrefutable belief systems had an overall poorer performance on the insight tasks. These findings suggest that frequent engagement with such beliefs may hinder the mental flexibility needed to adaptively restructure ideas. Given the importance of repetitive negative thinking across many psychopathologies, future research could potentially benefit from incorporating the degree of falsifiability as a novel transdiagnostic variable into contemporary etiological models.

Keywords: irrefutable reasoning, constraint relaxation, maladaptive perseveration

1 Introduction

1.1 Maladaptive Perseveration

Theoretical models have long delineated neurocognitive components, such as bottom-up limbic hyperactivity and top-down attenuated cognitive control, as key mechanisms underlying repetitive negative thoughts [19, 43, 31]. Beyond these dysfunction-based accounts, cognitive perseveration has also been linked to motivated reasoning [45, 7], where individuals are incentivized to process information in a biased manner that supports their pre-existing beliefs or desires. Relatedly, confirmation bias—the tendency to predominantly engage with information that confirms one's existing beliefs—has been shown to create self-reinforcing loops by amplifying congruent evidence and discounting disconfirming data [41, 51]. Prior work has also shown that perseverative cognition in negative mental states can mediate the development of chronic pathogenic conditions by inducing

*Corresponding author: stijn.vanseveren@gmail.com

a continuous psychophysiological burden [6]. Therefore, it remains vital to consider novel predictors when creating holistic models with incremental validity compared to traditional etiological models.

Since the falsifiability criterion was formalized as a key demarcation principle in rational inquiry [56], no experimental research has yet examined whether unfalsifiability can moderate cognitive perseveration. Some belief systems can be set within parameters that preclude disconfirmation (i.e., unfalsifiable), making them effectively impervious to contradictory evidence because such evidence can be rendered epistemically irrelevant to preserve internal coherence [56, 4]. Consequently, cognitive perseveration may occur not solely from dysfunctional neurocognitive systems, motivated reasoning, or confirmation bias, but also as a feature of the axiomatic framework on which the belief system rests.

1.2 De Facto Unfalsifiability

Two recurring types of irrefutable reasoning can achieve invulnerability against rational criticism and empirical disconfirmation [4]: (a) an *epistemic defense mechanism*, a structural (proposition-inherent) feature of a belief system allowing reframing of disconfirming accounts to maintain consistency; and (b) an *immunizing strategy*, a flexible (proposition-independent) feature applied externally, allowing disconfirming accounts to be dismissed, discredited, or reinterpreted to preserve validity.

Some modern belief systems are often criticized for relying heavily on irrefutable reasoning, including mono- and polytheistic doctrines [16], some postmodernist branches [55], psychoanalytic frameworks [2], alternative medicinal practices [23], conspiracy beliefs [3], Scientology [39], and the simulation hypothesis [65].

To make the construct experimentally tractable, a practical variant of irrefutability is introduced: *'A belief system, as a proposition, is "de facto" unfalsifiable (DFU) whenever there are proposition-inherent or proposition-independent epistemic features that effectively decrease its probability to be refuted'* [2].

1.3 Higher-order Priming Paradigm

To operationalize the impact of DFU belief systems on cognitive perseveration, a higher-order priming paradigm is employed where temporary adoption of DFU belief systems precedes measurement of insightful reasoning performance. Priming refers to the process by which exposure to a stimulus influences processing of and/or response to a subsequent stimulus, primarily via retained neural activation that makes related information or cognitive strategies more likely to shape subsequent response pathways [34, 13, 32].

Beyond priming of posterior perceptual processing, higher-order priming effects have been reported across multiple complex domains, including moral judgement [8], belief in free will or determinism [28], cooperation [21], risk perception [63], emotion regulation style [18], and pro-environmental attitudes [40]. Experimental exposure to unfalsifiable reasoning patterns has also been shown to increase ideological conviction and polarization [26].

1.4 Insightful Reasoning

Insight is often conceptualized as a transformative process where an individual abruptly transitions from not knowing how to solve a problem to knowing how to solve it [29, 58, 61]. According to Representational Change Theory, insight involves two restructuring mechanisms [42, 52]: (a) *constraint relaxation*, which entails inhibiting prior successful strategies to explore new feature dimensions within the current representation; and (b) *chunk decomposition*, where perceptual units are broken down to allow new configurations. Without these processes, an impasse period is reached during which further progress stalls.

The proposed link between insightful reasoning and DFU reasoning is the importance of re-evaluating the validity of the current representation. If habitual disregard of performance failure (contrary evidence) reduces the salience of that failure as a trigger to reconsider the representation, restructuring may be less likely to occur. This desensitization may generalize to problem-solving strategies, prolonging the impasse period during insight problems where constraint relaxation is crucial, thereby suppressing adaptive flexibility needed for representational change.

2 Methods and Materials

2.1 Design and Procedure

Sixty natively Dutch psychology students from Ghent University were recruited for a 55-minute PsychoPy experiment for which they could obtain course credit. To avoid uncontrolled between-participant differences and preserve statistical power, a within-subject design was used. The experiment comprised 24 trials, partially randomized into six blocks, where each block contained four trials of the same condition type. Each trial comprised three elements (Fig. 1):

Natural Reading Paradigm (1): Participants were instructed to temporarily simulate how "person A" would respond to critique by "person B". Participants had 90 seconds to read an in-first-person text of a cognitive schema and a critique by an antagonist, and to answer a multiple-choice question with three reaction options. The correct answer was always the one most aligned with how "person A" would respond [57].

An example DFU schema was the conspiracy belief that the absence of evidence for elitist corruption confirms its presence, under the assumption that powerful groups hide incriminating proof [50]. In contrast, an example of a refutable schema concerned the effectiveness of a fictional mRNA vaccine, phrased to make clear testable predictions.

Trials without a correct response on the multiple-choice question were treated as failures to adopt the schema. Overall accuracy on the multiple-choice questions was sufficient to test priming effects for most trials (M=80.11%; SD=18.27%). A logistic mixed-effects model showed that questions in DFU blocks were slightly easier (Δ=14.12%; p=0.00426). Reaction times did not significantly differ (Δ=1.76s; p=0.322). Random intercepts and slopes were included for participants, and random intercepts for task items.

To ensure that systematic differences between conditions only arose from falsifiability manipulation, a blind validation using a large language model was performed to rate potential confounders and the manipulation of interest. Thirty in-first-person texts were evaluated on valence, arousal, reading complexity, political neutrality, and degree of falsifiability [20, 59]. The 24 most suitable texts were selected such that the only significant difference between DFU and F texts was in falsifiability (d=-5.83; p=1.797e-12), while confounders showed small, non-significant effect sizes.

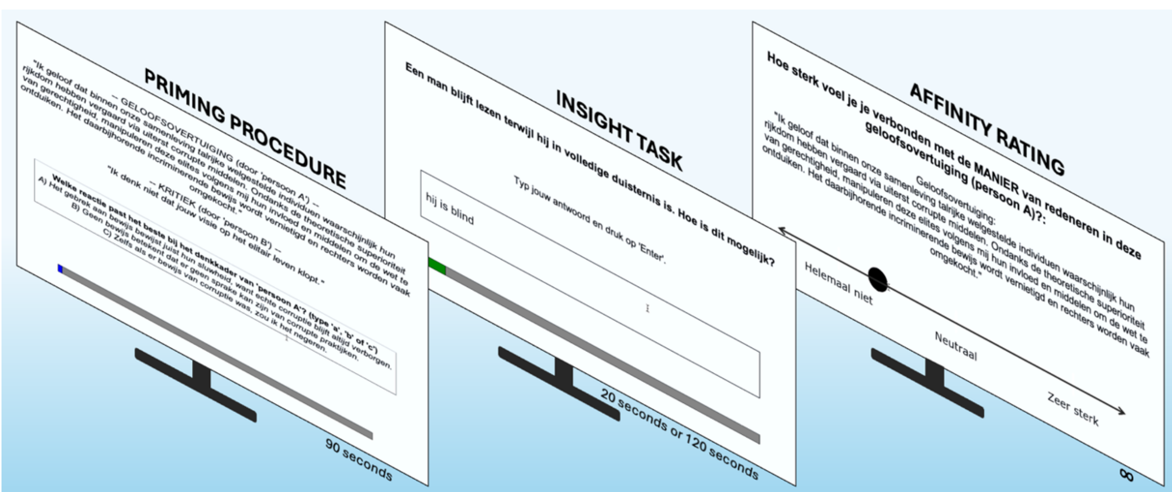
Insight Tasks (2): Each priming component was followed by an insight task. Classical insight tasks (120 s) and remote associates' tasks (RAT; 20 s) were used to balance internal validity and statistical power. Remaining time was shown via a visual analog.

A literature scan yielded 56 classical insight tasks; 44 were excluded using criteria on accuracy, response time, digital format feasibility, prior familiarity, and content type (verbal only), resulting in 12 classical tasks (Supplementary Table S1). A Dutch 22-item RAT variant was available [10]; the 10 most difficult items were removed to obtain 12 RAT items (Supplementary Table S2). Both task types were balanced across conditions.

Evidence supports applying Representational Change Theory to RAT [35]. Semantic search dynamics during RAT show patterns consistent with restructuring mechanisms [15]. The priming mechanism remains plausible because reduced propensity to shift feature spaces during DFU reasoning could reduce switching between semantic patches during RAT.

Affinity Ratings (3): Each trial ended with a 7-point slider rating of affinity toward "person A"'s cognitive schema (1=very low; 7=very strong). The schema text was displayed; no time constraint was imposed. After each block, a 10-second obligatory break reduced carry-over effects and mitigated attention/motivation depletion.

Figure 1: Exemplary sequence of a trial where its three constituents are illustrated



Note. Participants were instructed to read an in-first-person text of a belief system, which was either refutable or irrefutable; followed by a critique on that belief. Then, they had to answer a multiple-choice question where the correct answer is the option with the strongest aligning counter reaction. Afterwards, participants were given either a classical insight task or a remote associates' task. Finally, a 7-point slider was presented where the overall affinity towards the belief system could be scored.

2.2 Measures and Statistical Analysis

Classical and modern insight tasks were evaluated binarily (0=incorrect; 1=correct). Speed and accuracy were combined into a single balanced integration (BI) score designed to be relatively insensitive to speed–accuracy trade-offs and validated for within-subject designs [46, 47]. The formula was:

$$\mathrm{BI} = \frac{w \cdot z(\mathrm{Accuracy}) - (1-w) \cdot z(\mathrm{RT})}{\sqrt{w^2 + (1-w)^2}}$$

where w was initialized at 0.5, z denotes standardization of accuracy and reaction times, and the denominator normalizes the weight difference.

Analysis 1 (priming effect). Because priming relies on successful schema adoption, trials were excluded when multiple-choice responses were incorrect and no correct response had occurred earlier in that block. Four mixed-effects models were tested, crossed for unintegrated performance measures and task types, with maximal random effects. Significance was evaluated at Bonferroni-corrected $\alpha = 0.0125$. Two additional mixed-effects models were tested for both task types with fixed effects for condition, task type, and their interaction (threshold $\alpha = 0.025$). A final model integrated all task types and measures. Logistic mixed-effects models were used for accuracies; linear mixed-effects models were used for reaction times and BI scores.

Analysis 2 (DFU affinity association). Affinity ratings were averaged per participant for DFU and F schemata, yielding a DFU affinity index and an inverted F affinity index. The F index was inverted by subtracting F affinity scores from 8, enabling an overall affinity estimate using all schemata. Eight Pearson correlations tested associations between the two indices and aggregated accuracy/RT measures for classical and modern tasks (Bonferroni-corrected $\alpha = 0.00625$). Two additional Pearson correlations tested associations with overall BI score (Bonferroni-corrected $\alpha = 0.025$). Mixed-effects models with maximal random structure were fitted separately for accuracy, RT, and BI.

All analyses, except the mixed-effects models of the second analysis, were preregistered on OSF (`https://osf.io/xefsk/overview`). Mixed-effects analyses were conducted in R; correlation analyses and visualization in Python (`https://github.com/stvsever/RPEP_PROJECT`).

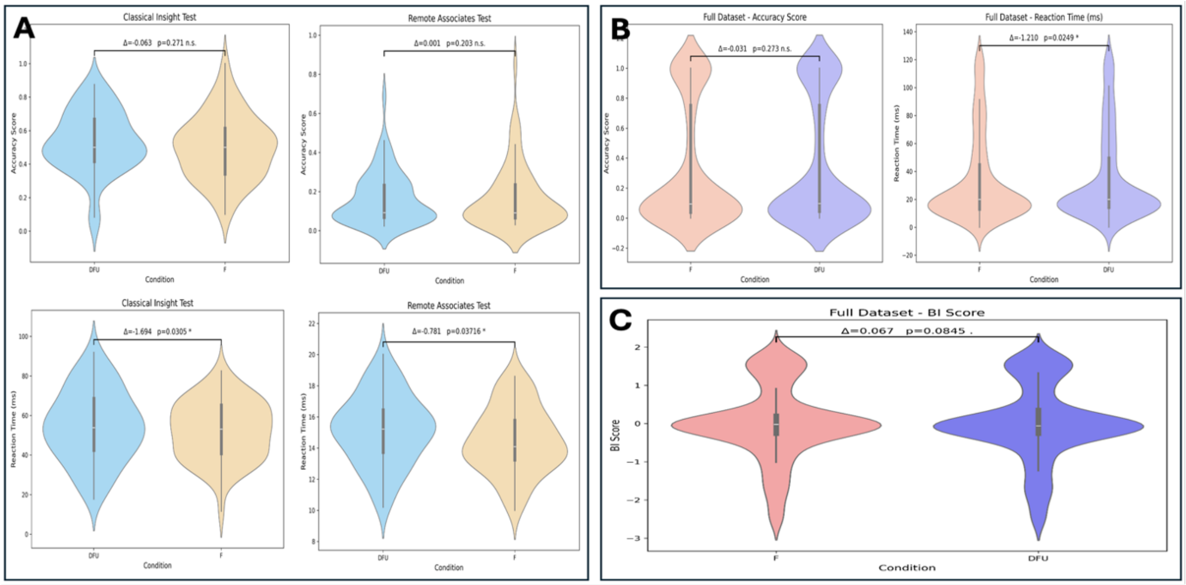
3 Results

3.1 Analysis 1: Higher-order Priming Effect

In total, 127 (8.82%) incorrect trials were excluded for which no priming effect could have occurred. One participant was excluded due to miscomprehension of the procedure. Mean accuracy for classical insight tasks was 54.25% (SD=49.76%), with no significant condition difference (Δ=0.063; p=0.271). Reaction times were higher in the DFU condition (Δ=1.69s; p=0.0305), with mean RT 53.63 seconds (SD=34.71s). For modern insight tasks, accuracy was lower (M=10.34%; SD=30.49%), with no accuracy difference (Δ=0.001%; p=0.203), whereas reaction times (M=14.79s; SD=46.48s) reached significance (Δ=0.781s; p=0.0372). No effects survived the Bonferroni-corrected threshold of 0.0125 (Fig. 2).

After combining both task types, the accuracy model again showed no significant condition difference (B=0.202; p=0.273). For reaction times, the model marginally survived the Bonferroni-corrected p-value of 0.025 (B=5.404; p=0.0249). The integrated BI score model did not reach p<0.05 (B=0.11; p=0.0845).

Figure 2: Results of higher-order priming effect analysis



Note. A) No significant differences in accuracy between conditions were found for both task types. However, reaction times were significantly higher in the DFU condition for both classical and modern insight tasks. B) A similar significance pattern was found for the full dataset where only reaction times significantly differed across conditions. C) Finally, an integrating approach to estimate the effect on overall insight performance resulted in marginal non-significance.

In a post-hoc analysis, four models were fitted to the reaction time distributions per item (ex-Gaussian, normal, chi-squared, polynomial), illustrating substantial variance across tasks (Supplementary Figure S1). Similar variance patterns were observed across participants and condition items (Supplementary Figure S2).

3.2 Analysis 2: Association between DFU Affinity and Insight Performance

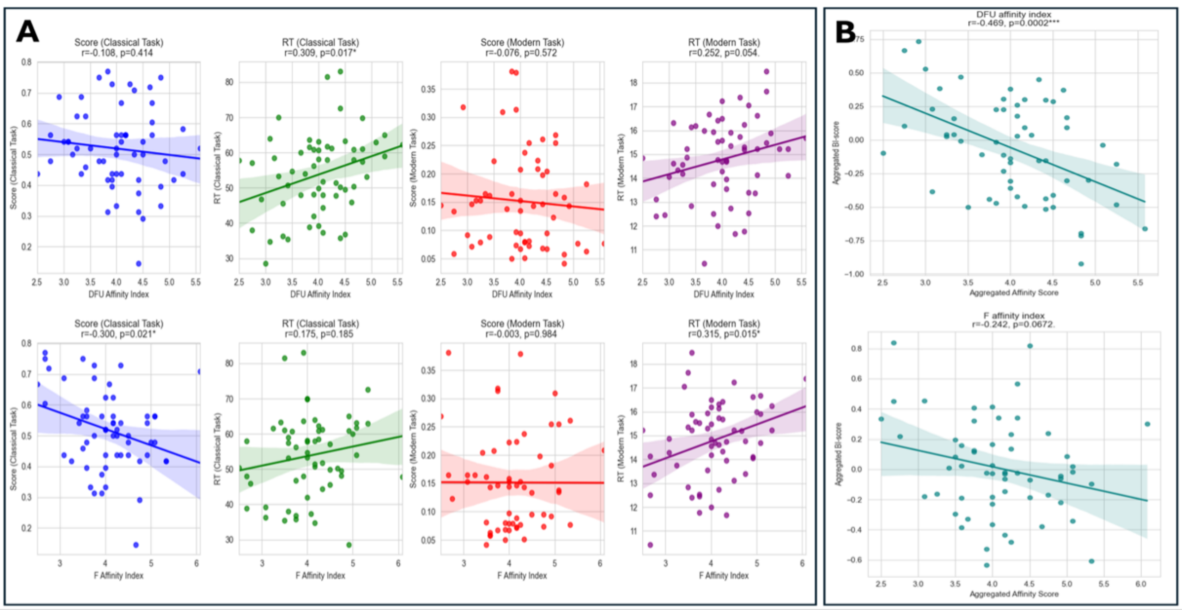
After aggregating affinity scores per participant and condition, no significant difference was found between DFU and inverted F affinity scores (Δ=-0.039; p=0.765). No correlation was found between the two indices (r=0.007; p=0.957). For the DFU affinity index, a significant association was observed only with reaction times for classical insight tasks (r=0.309; p=0.017). For the inverted F affinity index, associations were found with classical task accuracy (r=0.300; p=0.021) and modern task reaction time (r=0.315; p=0.015). None of these eight correlations survived Bonferroni correction ($\alpha = 0.00625$) (3).

After integrating and aggregating accuracy and RT across task types to compute an overall BI score per participant, a moderate negative correlation was observed between BI scores and DFU affinity (r=-0.469; p=0.0002). A weaker negative correlation was observed for the inverted F index, but this did not reach the Bonferroni-corrected threshold of 0.025 (r=-0.242; p=0.067).

Mixed-effects models were then fitted for accuracy, RT, and BI with random intercepts for participants, task items, and condition items. Self-perceived verbal intelligence was included as a fixed-effect covariate. Higher affinity was significantly associated with poorer performance for all three models (Supplementary Figure S3): accuracy (B=-0.0268; p=1.38e-08), reaction time (B=2.4276; p=2e-16), and BI score (B=-0.1146; p=3.14e-15).

A post-hoc factor analysis evaluated whether aggregation-by-averaging was appropriate for estimating latent disposition scores. Despite strong correlations (r>0.70) between averaged indices and factor-regression scores, extracted factors explained a small proportion of variance in affinity scores: 8% (DFU items), 10% (inverted F items), and 14% (combined) (Supplementary Figure S4).

Figure 3: Analysis results of association between affinity scores and insight performance



Note. A) For the DFU affinity index, a significant association was found only with reaction times on the classical insight tasks. In contrast, the inverted F affinity index showed significant correlations with both the accuracy of classical insight tasks and the reaction times of modern insight tasks. However, none of the eight Pearson correlation tests remained significant after applying Bonferroni correction. B) Overall insight performance (i.e., BI score) was moderately and significantly correlated with DFU affinity index, whereas its association with the inverted F index was weaker and fell just short of significance.

4 Discussion

This paper investigated how both acute (i.e., through higher-order priming) and presumed chronic exposure (i.e., through DFU affinity association) to irrefutable belief systems influence cognitive performance in domains where overcoming fixation and restructuring entrenched representations are essential for solving problems. The increasing digital proliferation of pseudoscientific ideas on social media platforms [24, 9, 37] raises pressing questions about how such exposure affects processing of and responses to conflicting information. The observed patterns are broadly consistent with the claim that engagement with unfalsifiable beliefs can negatively impact insightful reasoning. However, caution regarding the quality and limitations of the research materials is warranted.

Two types of insight tasks were used: classical verbal insight tasks and remote associates' items. Increasing trials per condition would have improved statistical power, but alternative tasks with shorter timeframes might have reduced internal validity relative to the proposed mechanism, namely that internal failure signals may become less likely to trigger representational updating. Among executive-function tasks that assess representational updating abilities, the Wisconsin Card Sorting Test aligns conceptually with the proposed mechanism [25], but once sorting rules are learned it primarily measures strategy switching under changing reinforcement rather than continued feature-based exploration. Classical insight tasks combined with RAT items therefore offered a pragmatic balance between internal validity and feasible study duration.

From theoretical and correlational perspectives, classical insight tasks and RAT share several properties: (a) both involve limited solution spaces and measure convergent creative thinking [33, 12]; (b) impasses arise from initially dominant but incorrect representations [5]; (c) constraint relaxation increases solution likelihood [15, 35, 52]; and (d) successful restructuring often coincides with a sudden "AHA!" experience that feels compelling and self-evident [14, 62].

4.1 Effect of Priming Procedure on Insight Performance

The predicted effect on overall insight performance (BI score) trended in the expected direction but did not reach statistical significance. When examining accuracy and reaction times separately, a consistent pattern emerged across task types: reaction times increased after exposure to rigid cognitive frameworks, consistent with a prolonged impasse period, while accuracy did not differ significantly between conditions. One interpretation is that accuracy may be relatively insensitive in insight problems, because participants tend to provide answers only when confident [62]. This could limit accuracy variation across conditions, particularly under longer time constraints that reduce guessing behavior.

The findings also motivate scrutiny of how "irrefutability" was operationalized. Although DFU items were validated to differ from refutable items specifically on falsifiability, subtle higher-order linguistic nuances may still have differed between conditions, potentially confounding downstream problem-solving processes. Another interpretation is that the higher correctness rate on multiple-choice questions in DFU blocks could have changed the allocation of cognitive resources, leading to subtle downstream changes in task engagement.

Methodologically, the brief priming paradigm (90 seconds reading/responding) may have limited depth of schema internalization. Future work could consider longer exposures or between-subject designs. Moreover, the sample comprised psychology students, who may engage with material more analytically than a general population sample. Replications in more ecologically valid contexts and more diverse samples are needed. Future work may also benefit from additional validation protocols or alternative designs to ensure participants meaningfully internalize the intended cognitive schemata.

4.2 Association between DFU affinity and Insight Performance

Affinity scores provided a second approach for examining how irrefutable belief systems relate to problem solving. The underlying assumption is that stronger affinities co-occur with stronger internalization and repeated exposure. Aggregation-by-averaging was used to estimate latent dispositions, and initial correlations partially aligned with the hypothesis: all correlations trended in predicted directions and several were below $p<0.05$, although none survived conservative Bonferroni correction.

When integrating speed and accuracy into overall BI scores, DFU affinity showed a moderate negative association with overall insight performance. These correlational findings were complemented by mixed-effects models using non-aggregated data with a covariate for self-perceived verbal intelligence [27]. This covariate was included because affinity toward pseudoscientific belief systems may correlate with verbal intelligence-related variables [60], and verbal insight itself correlates with verbal intelligence [17]. Given that self-perceived intelligence is only weakly correlated with objective intelligence [27], future studies should consider standardized verbal intelligence measures.

Range restriction may also have limited generalizability, as affinity scores did not occupy extreme ranges. For clinical research, in-depth qualitative case studies may be beneficial, especially when recruiting individuals with highly extreme belief endorsements is difficult.

The item construction and psychometric properties of the DFU/F items also warrant attention. Although large differences in falsifiability were achieved during validation, factor analyses suggested low explained variance for a single latent factor, indicating that better psychometrically validated items may be needed. Future work could use conventional item-evaluation frameworks such as Item Response Theory [66]. The lack of a strong negative correlation between DFU and inverted F indices further highlights potential context-dependence in endorsing specific belief systems, suggesting aggregation may be insufficient for latent trait estimation.

5 Implications and Future Directions

Both experimental and cross-sectional results point toward the possibility that adoption of irrefutable belief systems hinders the mental flexibility required to restructure ideas adaptively. Given the importance of repetitive negative thinking across psychopathologies [22, 67], future research could explore falsifiability as a novel transdiagnostic variable in contemporary etiological models. Increased clinical awareness may also support therapeutic processes that help individuals recognize and revise maladaptive cognitive frameworks [53].

Beyond clinical implications, it has been argued that the self-sealing nature of some conspiracy theories is a key ingredient in escalation toward more radical conspiratorial worldviews [3]. With the rise of pseudoscience across media platforms [37], platforms may benefit from incorporating falsifiability-related criteria into content flagging algorithms aimed at harm reduction [38, 30].

More research is needed on neurocognitive mechanisms behind impaired representational change under irrefutability. Electroencephalography approaches could test whether event-related potentials corroborate a reduced likelihood of adaptive restructuring when confronted with contrary evidence. A promising candidate is feedback-related negativity [36, 64], particularly in designs requiring responses to conflicting information. Prior work indicates that incongruent evaluative feedback can elicit frontal processing that drives behavioral updating [54]. The present account would predict attenuated neural responses and reduced behavioral adjustment in individuals with stronger DFU affinity when confronted with conflicting self-relevant information.

Functional magnetic resonance imaging studies employing representational similarity analysis [44, 49] could test whether poorer working-memory updating is observed under higher DFU affinity. Within a cascade-of-control framing [1], performance monitoring and representational maintenance involve dorsal anterior cingulate and dorsolateral prefrontal systems interacting with posterior association regions [48, 11]. The probability of triggering representational change after detected failure could plausibly vary by DFU affinity. Stronger autocorrelation in multivariate time series representing working-memory content might indicate reduced exploration of alternative representations under elevated DFU affinity.

Finally, item-level and factor-analytic findings indicate that current DFU items may not optimally estimate a general disposition toward irrefutable reasoning. Future studies should prioritize psychometrically validated items, and might benefit from qualitative comparative case studies that isolate idiosyncratic belief systems in context.

6 Conclusion

The results provide preliminary evidence consistent with the hypothesis that irrefutable belief systems can impair insightful reasoning by delaying the point at which adaptive representational change is triggered. This effect appeared under both acute priming and chronic exposure proxies, particularly in increased reaction times. Future research should use alternative designs and psychometrically validated materials to corroborate and refine these findings.

References

[1] Banich, M. T. (2019). The Stroop effect occurs at multiple points along a cascade of control: Evidence from cognitive neuroscience approaches. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.02164

[2] Boudry, M., Braeckman, J., & Buekens, F. (2007). *De naakte keizers van de psychoanalyse: de immunisatiestrategieën van een pseudowetenschap*. https://lib.ugent.be/nl/catalog/rug01:001396449

[3] Boudry, M. (2022). Why we should be suspicious of conspiracy theories: A novel demarcation problem. *Episteme, 20*(3), 611–631. https://doi.org/10.1017/epi.2022.3426

[4] Boudry, M., & Braeckman, J. (2010). Immunizing strategies and epistemic defense mechanisms. *Philosophia, 39*(1), 145–161. https://doi.org/10.1007/s11406-010-9254-9

[5] Bowden, E. M., & Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers, 35*(4), 634–639. https://doi.org/10.3758/BF03195543

[6] Brosschot, J. F., Gerin, W., & Thayer, J. F. (2006). The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. *Journal of Psychosomatic Research, 60*(2), 113–124. https://doi.org/10.1016/j.jpsychores.2005.06.074

[7] Caddick, Z. A., & Feist, G. J. (2021). When beliefs and evidence collide: Psychological and ideological predictors of motivated reasoning about climate change. *Thinking & Reasoning, 28*(3), 428–464. https://doi.org/10.1080/13546783.2021.1994009

[8] Cameron, C. D., Payne, B. K., Sinnott-Armstrong, W., Scheffer, J. A., & Inzlicht, M. (2017). Implicit moral evaluations: A multinomial modeling approach. *Cognition, 158*, 224–241. https://doi.org/10.1016/j.cognition.2016.10.013

[9] Chavda, V. P., Sonak, S. S., Munshi, N. K., & Dhamade, P. N. (2022). Pseudoscience and fraudulent products for COVID-19 management. *Environmental Science and Pollution Research, 29*(42), 62887–62912. https://doi.org/10.1007/s11356-022-21967-4

[10] Chermahini, A. S., Hickendorff, M., & Hommel, B. (2012). Development and validity of a Dutch version of the remote associates task: An item-response theory approach. *Thinking Skills and Creativity, 7*(3), 177–186. https://doi.org/10.1016/j.tsc.2012.02.003

[11] Chiang, T., Lu, R., Hsieh, S., Chang, Y., & Yang, Y. (2014). Stimulation in the dorsolateral prefrontal cortex changes subjective evaluation of percepts. *PLoS ONE, 9*(9), e106943. https://doi.org/10.1371/journal.pone.0106943

[12] Colzato, L. S., Ozturk, A., & Hommel, B. (2012). Meditate to create: The impact of focused-attention and open-monitoring training on convergent and divergent thinking. *Frontiers in Psychology, 3*. https://doi.org/10.3389/fpsyg.2012.00116

[13] Dai, W., Yang, T., White, B. X., Palmer, R., Sanders, E. K., McDonald, J. A., Leung, M., & Albarracín, D. (2023). Priming behavior: A meta-analysis of the effects of behavioral and

nonbehavioral primes on overt behavioral outcomes. *Psychological Bulletin, 149*(1–2), 67–98. https://doi.org/10.1037/bul0000374

[14] Danek, A. H., Williams, J., & Wiley, J. (2018). Closing the gap: Connecting sudden representational change to the subjective AHA! experience in insightful problem solving. *Psychological Research, 84*(1), 111–119. https://doi.org/10.1007/s00426-018-0977-8

[15] Davelaar, E. J. (2015). Semantic search in the remote associates test. *Topics in Cognitive Science, 7*(3), 494–512. https://doi.org/10.1111/tops.12146

[16] Dawkins, R. (2006). *The God Delusion*. Boston, MA: Houghton Mifflin. ISBN 0-618-68000-4.

[17] Dechaume, M., Mercier, M., Feybesse, C., Lubart, T., Chouvelon, G., Kermarrec, S., & Tordjmann, S. (2024). Is intelligence necessary and sufficient for creativity? An analysis of convergent and divergent *Elsevier, 117*, 1041-6080. https://scholar.google.com/scholar?q=dechaume+2024+verbal+insight+intelligence

[18] Dewitte, M. (2011). Adult attachment and attentional inhibition of interpersonal stimuli. *Cognition & Emotion, 25*(4), 612–625. https://doi.org/10.1080/02699931.2010.508683

[19] Disner, S. G., Beevers, C. G., Haigh, E. A., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience, 12*(8), 467–477. https://doi.org/10.1038/nrn3027

[20] Dreisbach, G., & Fischer, R. (2012). Conflicts as aversive signals. *PsycEXTRA Dataset*. https://doi.org/10.1037/e502412013-956

[21] Drouvelis, M., Metcalfe, R., & Powdthavee, N. (2015). Can priming cooperation increase public good contributions? *Theory and Decision, 79*(3), 479–492. https://doi.org/10.1007/s11238-015-9481-4

[22] Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic process. *International Journal of Cognitive Therapy, 1*(3), 192–205. https://doi.org/10.1521/ijct.2008.1.3.192

[23] Ernst, E., & Singh, S. (2009). *Trick or Treatment? Alternative Medicine on Trial*. London: Transworld Publishers.

[24] Faddoul, M., Chaslot, G., & Farid, H. (2020). A longitudinal analysis of YouTube's promotion of conspiracy videos. *arXiv*. https://doi.org/10.48550/arXiv.2003.03318

[25] Friedman, N. P., & Robbins, T. W. (2021). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology, 47*(1), 72–89. https://doi.org/10.1038/s41386-021-01132-0

[26] Friesen, J. P., Campbell, T. H., & Kay, A. C. (2015). The psychological advantage of unfalsifiability: The appeal of untestable religious and political ideologies. *Journal of Personality and Social Psychology, 108*(3), 515–529. https://doi.org/10.1037/pspp0000018

[27] Furnham, A., & Grover, S. (2020). Correlates of Self-Estimated Intelligence. *Journal of Intelligence*, 8(1), 6. https://doi.org/10.3390/jintelligence8010006

[28] Genschow, O., Cracco, E., Schneider, J., Protzko, J., Wisniewski, D., Brass, M., & Schooler, J. (2021). *PsyArXiv preprint* Manipulating belief in free will and its downstream consequences: A meta-analysis. https://doi.org/10.31234/osf.io/quwgr

[29] Gilhooly, K., & Murphy, P. (2005). Differentiating insight from non-insight problems. *Thinking & Reasoning, 11*(3), 279–302. https://doi.org/10.1080/13546780442000187

[30] Gomes, A. B., & Sultan, A. (2024). Problematizing content moderation by social media platforms and its impact on digital harm reduction. *Harm Reduction Journal, 21*(1). https://doi.org/10.1186/s12954-024-01104-9

[31] Hallion, L. S., Wright, A. G., Joormann, J., Kusmierski, S. N., Coutanche, M. N., & Caulfield, M. K. (2022). A five-factor model of perseverative thought. *Journal of Psychopathology and Clinical Science, 131*(3), 235–252. https://doi.org/10.1037/abn0000737

[32] Hao, J., Plangger, K., & West, D. (2024). Conceptualizing sustainable consumption priming: A scoping review. *Psychology & Marketing, 41*(11), 2772–2788. https://doi.org/10.1002/mar.22083

[33] Hommel, B., Colzato, L. S., Fischer, R., & Christoffels, I. K. (2011). Bilingualism and Creativity: Benefits in Convergent Thinking Come with Losses in Divergent Thinking. *Frontiers in Psychology*, 2. https://doi.org/10.3389/fpsyg.2011.00273

[34] Horner, A., & Henson, R. (2008). Priming, response learning and repetition suppression. *Neuropsychologia, 46*(7), 1979–1991. https://doi.org/10.1016/j.neuropsychologia.2008.01.018

[35] Huang, P., Liu, C., & Chen, H. (2019). Examining the applicability of representational change theory for remote associates problem-solving with eye movement evidence. *Thinking Skills and Creativity, 31*, 198–208. https://doi.org/10.1016/j.tsc.2018.12.001

[36] Huang, Y., & Yu, R. (2014). The feedback-related negativity reflects more or less prediction error in appetitive and aversive conditions. *Frontiers in Neuroscience, 8*. https://doi.org/10.3389/fnins.2014.00108

[37] C. Impey (2024) DEVELOPING TECHNOLOGY TOOLS TO COMBAT FAKE SCIENCE, *EDULEARN24*, Proceedings, pp. 493-500. https://scholar.google.com/scholar?q=impey+2024+pseudoscience+social+media

[38] Katsaros, M., Kim, J., & Tyler, T. (2023). Online content moderation: Does justice need a human face? *International Journal of Human–Computer Interaction, 40*(1), 66–77. https://doi.org/10.1080/10447318.2023.2210879

[39] Kent, S. (1999). Creation of 'Religious' Scientology. *Religious Studies and Theology, 18*(2), 97–126. https://doi.org/10.1558/rsth.v18i2.97

[40] Kim, E. J., Tanford, S., & Book, L. A. (2020). The effect of priming and customer reviews on sustainable travel behaviors. *Journal of Travel Research, 60*(1), 86–101. https://doi.org/10.1177/0047287519894069

[41] Klayman, J. (1995). Varieties of confirmation bias. *Psychology of Learning and Motivation*, 385–418. https://doi.org/10.1016/S0079-7421(08)60315-1

[42] Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(6), 1534–1555. https://doi.org/10.1037/0278-7393.25.6.1534

[43] Koster, E. H., De Lissnyder, E., Derakshan, N., & De Raedt, R. (2011). Understanding depressive rumination from a cognitive science perspective: The impaired disengagement hypothesis. *Clinical Psychology Review, 31*(1), 138–145. https://doi.org/10.1016/j.cpr.2010.08.005

[44] Kriegeskorte, N. (2008). Representational similarity analysis—connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*. https://doi.org/10.3389/neuro.06.004.2008

[45] Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

[46] Liesefeld, H. R., & Janczyk, M. (2018). Combining speed and accuracy to control for speed-accuracy trade-offs. *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

[47] Liesefeld, H. R., & Janczyk, M. (2022). Correction to: Same same but different: Subtle but consequential differences between two measures to linearly integrate speed and accuracy (LISAS vs. BIS). *Behavior Research Methods, 55*(3), 1511–1511. https://doi.org/10.3758/s13428-022-01963-9

[48] MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*(5472), 1835–1838. https://doi.org/10.1126/science.288.5472.1835

[49] Matheson, H. E., Kenett, Y. N., Gerver, C., & Beaty, R. E. (2023). Representing creative thought: A representational similarity analysis of creative idea generation and evaluation. *Neuropsychologia, 187*, 108587. https://doi.org/10.1016/j.neuropsychologia.2023.108587

[50] Mehl, S., Rief, W., Soll, D., & Pytlik, N. (2025). Populist attitudes and belief in conspiracy theories: Anti-elitist attitudes and the preference for unrestricted popular sovereignty reduce the positive impact of an analytical thinking style on conspiracy beliefs. *BMC Research Notes, 18*(1). https://doi.org/10.1186/s13104-025-07136-z

[51] Modgil, S., Singh, R. K., Gupta, S., et al. (2024). A confirmation bias view on social media induced polarisation during COVID-19. *Information Systems Frontiers, 26*, 417–441. https://doi.org/10.1007/s10796-021-10222-9

[52] Osuna-Mascaró, A. J., & Auersperg, A. M. (2021). Current understanding of the "Insight" phenomenon across disciplines. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.791398

[53] Overholser, J. C., & Beale, E. (2023). The art and science behind Socratic questioning and guided discovery: A research review. *Psychotherapy Research, 33*(7), 946–956. https://doi.org/10.1080/10503307.2023.2183154

[54] Peters, A., Witte, J., Helming, H., Moeck, R., Straube, T., & Schindler, S. (2025). How positive and negative feedback following real interactions changes subsequent sender ratings. *Scientific Reports, 15*(1). https://doi.org/10.1038/s41598-025-91750-1

[55] Pluckrose, H., & Lindsay, J. (2020). *Cynical theories : how activist scholarship made everything about race, gender, and identity-and why this harms everybody*, Pitchstone Publishing. https://scholar.google.com/scholar?q=Pluckrose+Lindsay+Cynical+Theories+year

[56] Popper, K. (1934/1935). *Logik der Forschung: Zur Erkenntnistheorie der modernen Naturwissenschaft.* Julius Springer.

[57] Reinecke, A., Rinck, M., Becker, E. S., & Hoyer, J. (2013). Cognitive-behavior therapy resolves implicit fear associations in generalized anxiety disorder. *Behaviour Research and Therapy, 51*(1), 15–23. https://doi.org/10.1016/j.brat.2012.10.004

[58] Sheth, B. R., Sandkühler, S., & Bhattacharya, J. (2009). Posterior beta and anterior gamma oscillations predict cognitive insight. *Journal of Cognitive Neuroscience, 21*(7), 1269–1279. https://doi.org/10.1162/jocn.2009.21069

[59] Shenhav, A., Botvinick, M., & Cohen, J. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron, 79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

[60] Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science, 29*(8), 1358–1369. https://doi.org/10.1177/0956797618774253

[61] Stuyck, H., Aben, B., Cleeremans, A., & Van den Bussche, E. (2021). The AHA! moment: Is insight a different form of problem solving? *Consciousness and Cognition, 90*, 103055. https://doi.org/10.1016/j.concog.2020.103055

[62] Stuyck H, Cleeremans A, & Van den Bussche E. Aha! under pressure: The Aha! experience is not constrained by cognitive load. Cognition. (2022) doi:10.1016/j.cognition.2021.104946.

[63] Wadhera, T., & Kakkar, D. (2020). Modeling risk perception using independent and social learning: Application to individuals with autism spectrum disorder. *The Journal of Mathematical Sociology, 45*(4), 223–245. https://doi.org/10.1080/0022250X.2020.1774877

[64] Wang, Y., Cheung, H., Yee, L. T., & Tse, C. (2020). Feedback-related negativity (FRN) and theta oscillations: Different feedback signals for non-conform and conform decisions. *Biological Psychology, 153*, 107880. https://doi.org/10.1016/j.biopsycho.2020.107880

[65] Wolpert, D. (2025). What computer science has to say about the simulation hypothesis. Journal of Physics: Complexity. https://doi.org/10.1088/2632-072x/ae1e50

[66] Yigiter, M. S., & Boduroglu, E. (2024). Item Response Theory assumptions: A comprehensive review of studies with document analysis. *International Journal of Educational Studies and Policy, 5*(2). https://doi.org/10.5281/zenodo.14016086

[67] Zagaria, A., Ballesio, A., Vacca, M., & Lombardo, C. (2023). Repetitive negative thinking as a central node between psychopathological domains: A network analysis. *International Journal of Cognitive Therapy, 16*(2), 143–160. https://doi.org/10.1007/s41811-023-00162-4

Supplementary Materials

Supplementary Table S1

Table S1: Classical insight items used in the experiment

Task ID	Classical Insight Problem
1	"A window cleaner falls from a 12-meter ladder onto a concrete surface, but is not injured. How is that possible?"
2	"What has cities without houses, forests without trees, and rivers without water?"
3	"What can be broken without ever being touched or seen?"
4	"What occurs once in a minute, twice in a moment, but never in a thousand years?"
5	"What travels around the world but stays in one corner?"
6	"A man keeps reading while he is in complete darkness. How is this possible?"
7	"How can someone walk on the surface of a lake without sinking and without using any aids?"
8	"Ruben is taking part in a running race on Friday. He runs faster than Mark, who trains with him every Monday, Wednesday, and Friday. Despite this, Maria's coach, He is slower than the coach, but faster than Ruben. Arrange the FOUR runners from fastest to slowest using '>' to separate them."
9	"An antique coin dealer received an offer to purchase a beautiful bronze coin. The coin featured an emperor's head on one side and the date '544 BC' on the other. The dealer examined the coin, but instead of buying it, he called the police to arrest the man. Why did the dealer suspect that the coin was fake?"
10	"Using only a 7-minute hourglass and an 11-minute hourglass, how can you time exactly 15 minutes to cook an egg?"
11	"What can go up and come down without ever moving?"
12	"I am not alive, but I grow; I have no lungs, but I breathe; I have no mouth, yet water kills me. What am I?"

Note. Out of 56 identified classical insight tasks, 44 were excluded based on several criteria: unsuitable accuracy levels, excessive response times, digital formatting issues, potential prior exposure, and non-verbal content. This selection procedure resulted in a refined set of 12 verbal insight tasks appropriate for this study's framework. The correct answers to the classical insight tasks can be provided by the author upon request.

Supplementary Table S2

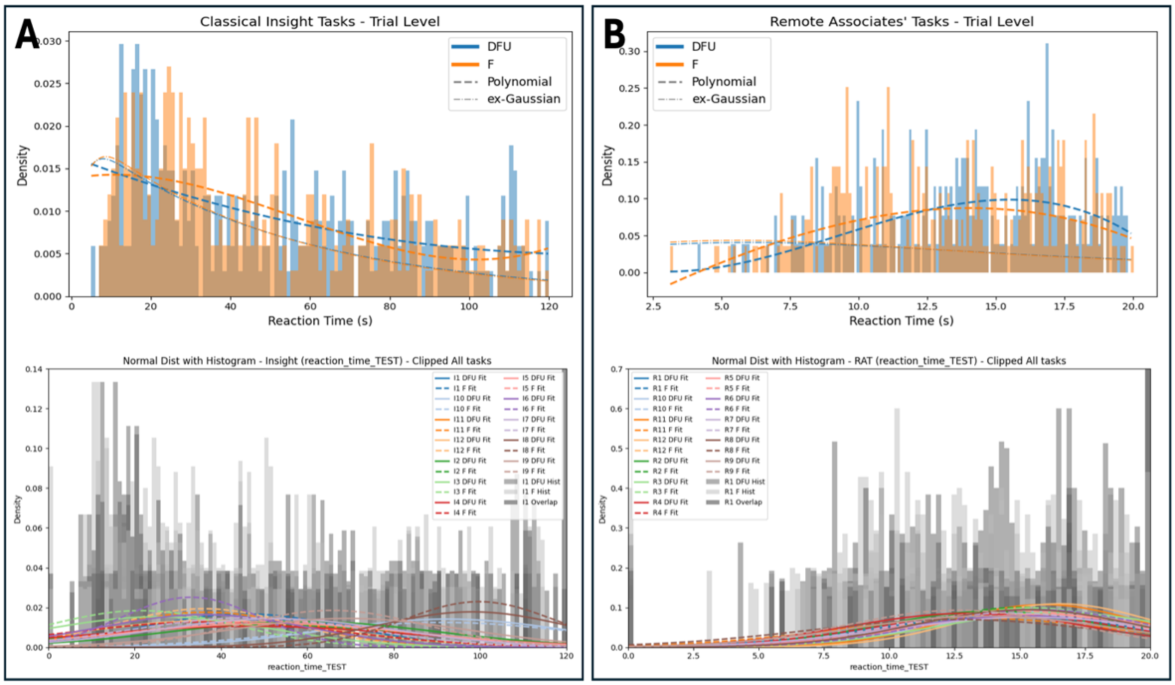
Table S2: Cue words for Dutch Remote Associates' Test (RAT) items used in the experiment

Task ID	Cue Words
1	["Bar", "jurk", "glas"]
2	["Kaas", "land", "huis"]
3	["Vlokken", "ketting", "pet"]
4	["Val", "meloen", "lelie"]
5	["Vig", "mijn", "geel"]
6	["Achter", "kruk", "mat"]
7	["Worm", "kast", "tegger"]
8	["Water", "schoorsteen", "lucht"]
9	["Trommel", "beleg", "mes"]
10	["Hond", "druk", "band"]
11	["Controle", "plaats", "gewicht"]
12	["Goot", "kool", "bak"]

Note. Another approach to estimating convergent creativity is to use remote associates' items. A Dutch 22-item version was developed using Item Response Theory. Twelve items were selected by excluding the 10 most difficult. Each item requires identifying a single target word that matches three seemingly unrelated cue words. Correct answers can be provided by the author upon request.

Supplementary Figure S1

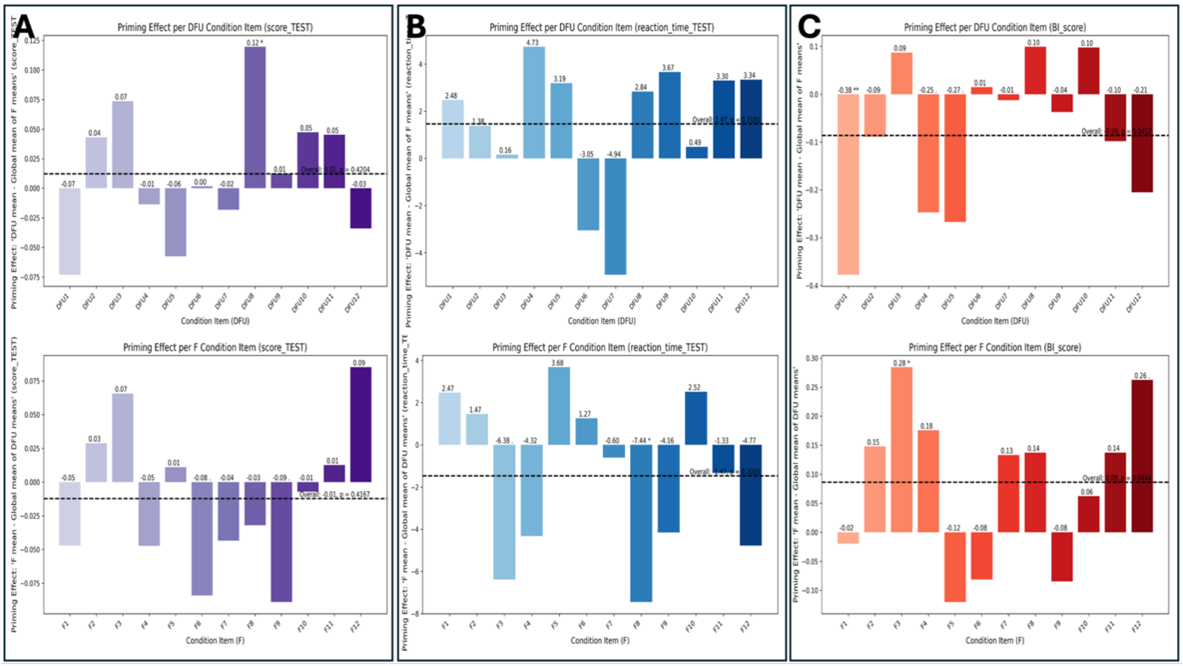
Figure S1: Results of reaction time distributions fitted with polynomial, ex-Gaussian and normal models



Note. A) For the classical insight tasks, a higher density for the DFU condition is seen for higher reaction times compared to the F condition (upper plot). When fitting the preprocessed reaction time data with 2-parameter Gaussian models separately for experimental conditions and task IDs, the substantial between-tasks performance can be illustrated (lower plot). B) A noticeably similar pattern is present for the RAT data where higher densities are found for the DFU condition for higher reaction times compared to the F condition.

Supplementary Figure S2

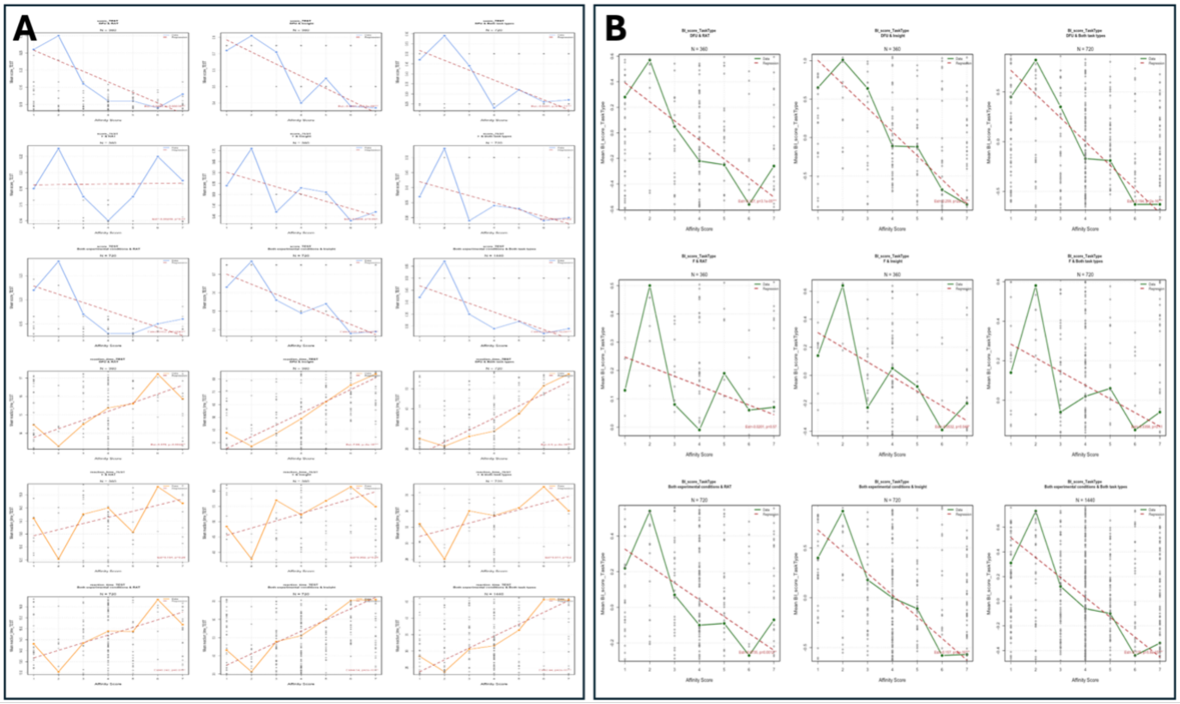
Figure S2: Item-level analysis illustrates variation in priming effect across different condition IDs



Note. The between-conditionIDs variation is shown in the upper and lower plots for the DFU condition and F condition, respectively. Each value was computed by taking the mean of the respective condition ID values and subtracting it by the global mean of all 12 condition ID values from the other experimental condition. An unpaired Welch's t-test was run for each separate higher-order priming effect across all three performance measures. A) For both experimental conditions, the respective mean accuracies non-significantly moved towards the direction that did not align with the hypothesis. Only 10 out of 24 items showed a pattern consistent with the hypothesis. B) For the reaction times, the predicted pattern occurred for both experimental conditions. However, two DFU items and five F items trended towards the opposite direction. C) For the BI scores, most condition IDs gravitated towards the predicted directions. On average, this was the case for the condition IDs in both experimental conditions.

Supplementary Figure S3

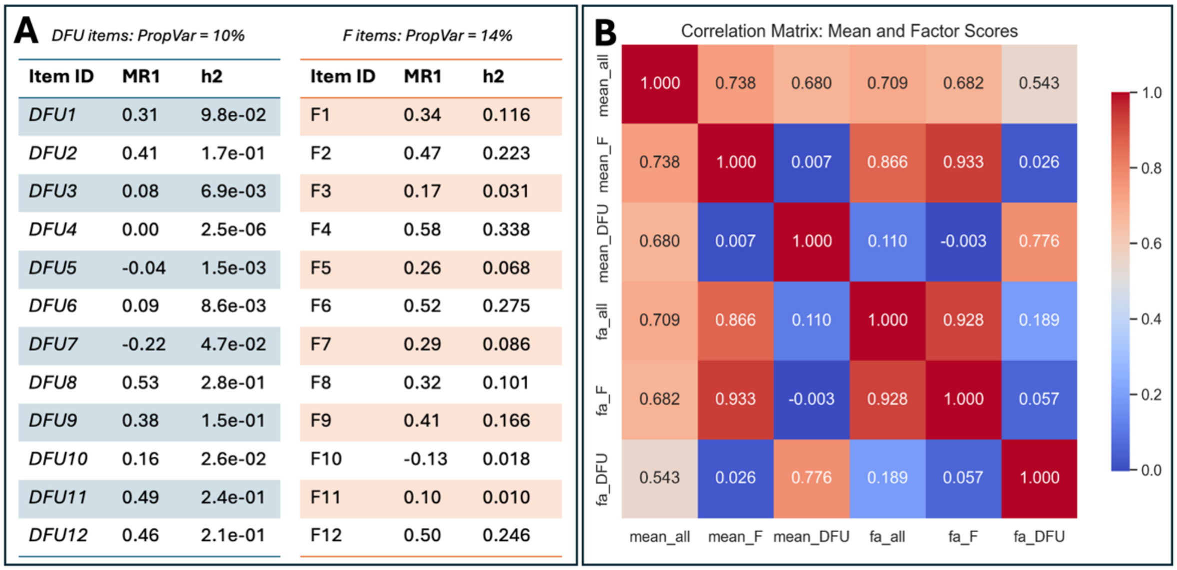
Figure S3: Post-hoc analysis with mixed-effects models with non-aggregated data for all three measures



Note. Using maximal random effects structure, the statistical association between the non-aggregated affinity scores and insight performance was estimated. In each model, self-perceived verbal intelligence was included as a fixed effect. Rows vary in experimental condition (1: DFU, 2: F, 3: both); columns vary in task type (1: RAT, 2: classical, 3: both). A) For each unit increase in affinity towards unfalsifiability, the accuracy (upper blue plots) decreased by 2.68% (B=0.0268; p=1.38e-08). For each unit increase, the reaction time (bottom orange plots) increased by 2.43 seconds (B=2.427; p=2e-16). B) The overall insight performance decreased by approximately one tenth of a standard deviation for each unit increase in DFU affinity (B=-0.1146; p=3.14e-15).

Supplementary Figure S4

Figure S4: Results of post-hoc factor analysis to evaluate latent disposition estimation techniques



Item ID	MR1	h2
DFU1	0.31	9.8e-02
DFU2	0.41	1.7e-01
DFU3	0.08	6.9e-03
DFU4	0.00	2.5e-06
DFU5	-0.04	1.5e-03
DFU6	0.09	8.6e-03
DFU7	-0.22	4.7e-02
DFU8	0.53	2.8e-01
DFU9	0.38	1.5e-01
DFU10	0.16	2.6e-02
DFU11	0.49	2.4e-01
DFU12	0.46	2.1e-01

Item ID	MR1	h2
F1	0.34	0.116
F2	0.47	0.223
F3	0.17	0.031
F4	0.58	0.338
F5	0.26	0.068
F6	0.52	0.275
F7	0.29	0.086
F8	0.32	0.101
F9	0.41	0.166
F10	-0.13	0.018
F11	0.10	0.010
F12	0.50	0.246

	mean_all	mean_F	mean_DFU	fa_all	fa_F	fa_DFU
mean_all	1.000	0.738	0.680	0.709	0.682	0.543
mean_F	0.738	1.000	0.007	0.866	0.933	0.026
mean_DFU	0.680	0.007	1.000	0.110	-0.003	0.776
fa_all	0.709	0.866	0.110	1.000	0.928	0.189
fa_F	0.682	0.933	-0.003	0.928	1.000	0.057
fa_DFU	0.543	0.026	0.776	0.189	0.057	1.000

Note. A) Factor loadings and communalities were computed three times with one latent factor. The separate results for the 12 DFU items (left table) and 12 F items (right table) are displayed. The proportion of explained variance was considerably low for both DFU items and F items: 10% and 14%, respectively. B) A pairwise correlation matrix was constructed by computing all correlations between the three factor analysis regression scores and the three mean scores for each participant. Although a strong correlation between the two latent disposition estimation methods was found (0.709>r>0.933) for all three datasets (i.e., 12 DFU items, 12F items, and 24 DFU+F items), future research could benefit from using better condition items to derive a single latent disposition score for each participant.