

Contents

The DBN Trader Rating: An ELO-Inspired Composite Metric for Behavioral and Outcome-Based Trader Development Assessment	1
Abstract	1
1. Introduction and Motivation	2
2. System Design Philosophy	2
3. Rating Computation	3
4. Asymmetric Penalty Rationale	5
5. The Density-Rating Bridge	6
6. Regression Detection and Recovery	7
7. Marketplace Integration	8
8. Volatility Dampening and Stability	9
9. Validation Methodology	10
10. Limitations and Open Questions	10
11. Invitation for Peer Review	11

The DBN Trader Rating: An ELO-Inspired Composite Metric for Behavioral and Outcome-Based Trader Development Assessment

Drive By Numbers (DBN) Technical White Paper Version 1.0 - April 2026

Abstract

Traditional trader evaluation metrics such as win rate, profit factor, and net profit-and-loss reduce a complex, multi-dimensional performance profile to a single outcome axis. These measures fail to distinguish between a disciplined trader experiencing a normal losing streak and an undisciplined trader on a lucky winning streak. This paper introduces the DBN Trader Rating, an ELO-inspired composite metric that fuses outcome signals (win/loss) with behavioral signals (revenge trading, stop-loss discipline, plan adherence) into a single scalar rating that evolves with every closed trade. The system begins each trader at a base rating of 1200 and applies asymmetric per-trade deltas that penalize process failures more heavily than outcome failures. At a threshold of 1600, the platform’s interface automatically transitions from a full-density pedagogical layout to a minimal, expert-oriented layout, creating a direct feedback loop between measured competence and information architecture. A three-trigger regression detection mechanism monitors for behavioral deterioration and automatically reverses the interface transition when sustained decline is detected. The Rating further serves as a trust signal in the platform’s strategy marketplace, gating publication privileges and weighting search rankings. This paper details the mathematical formulation, the pedagogical reasoning behind the asymmetric penalty structure, the density-rating bridge, the regression detection algorithm, marketplace integration, stability mechanisms, preliminary validation methodology, and acknowledged limitations.

1. Introduction and Motivation

The evaluation of trader performance is, at first glance, a solved problem. Financial markets produce unambiguous numerical outcomes on every closed position: a trader either made money or lost it. From this elementary binary, the industry has constructed a familiar apparatus of derived metrics - win rate, average risk-to-reward ratio, Sharpe ratio, Sortino ratio, maximum drawdown, profit factor, expectancy. These metrics are indispensable for strategy evaluation, but they share a critical blind spot: they measure what happened, not how it happened.

Consider two traders who each close one hundred trades over a month, both achieving a 55% win rate with a profit factor of 1.4. By every conventional metric, they are indistinguishable. Yet Trader A follows a pre-trade checklist on every entry, respects stop-loss placements, and spaces trades to avoid emotional decision-making after losses. Trader B enters impulsively after losses, moves stop-losses to avoid planned losses, and frequently deviates from the trading plan. Any experienced trading psychologist would predict dramatically different long-term trajectories. Trader A's process is sustainable; Trader B's is not.

The DBN Trader Rating addresses this gap by constructing a composite scalar that blends outcome signals with behavioral signals in a single evolving number. The design draws inspiration from the Elo rating system used in competitive chess, where a player's rating is updated after each game based on both the result and the context in which it occurred. In the DBN system, each closed trade functions analogously to a completed chess game: the result matters, but the behavioral context of the trade - whether it was a revenge entry, whether the stop-loss was moved, whether the trader followed a pre-trade plan - modifies the rating update independently of the profit-and-loss outcome.

The choice of an Elo-like framework is deliberate. Elo systems produce a single number that is easily communicated, compared, and tracked, avoiding the cognitive burden of monitoring a dashboard of partially correlated metrics. They are update-efficient: each new data point (trade) requires only a constant-time computation to incorporate, which is essential for a system operating in real-time on a client-side application with no backend computation budget.

2. System Design Philosophy

The Trader Rating rests on a conceptual distinction between two categories of signal that a closed trade can emit: outcome signals and behavioral signals. Outcome signals are the conventional profit-and-loss data that every trading platform captures. Behavioral signals are metadata about how the trade was conducted - metadata that is typically discarded or, at best, logged in a journal that the trader may or may not review.

The design philosophy of the Rating system treats behavioral signals as leading indicators and outcome signals as lagging indicators. This framing inverts the conventional priority. In most trading platforms, a winning trade is unambiguously positive and a losing trade is unambiguously negative. In the DBN Rating, a winning trade that was entered as a revenge trade after a loss, with a moved stop-loss and no pre-trade checklist, produces a net negative rating delta despite the positive P&L. Conversely, a losing trade that was entered with high plan adherence, an unmoved stop-loss, and proper spacing from prior trades produces only a modest negative delta. The system explicitly encodes the proposition that a good process with a bad outcome is preferable to a bad process with a good outcome, because the former is sustainable and the latter is not.

This design choice positions the Rating as a “character rating” rather than a “performance rating.” The term is borrowed from insurance underwriting, where an applicant’s character rating reflects their behavioral reliability independent of any single claim event. A trader’s Rating in the DBN system reflects their behavioral reliability independent of any single trade outcome. Over a sufficient number of trades, behavioral reliability and outcome quality converge – the Rating simply reaches convergence faster by incorporating both signal types simultaneously rather than waiting for behavioral deficiencies to manifest as outcome deficiencies.

The system is intentionally opinionated. It embeds a normative model of what “good trading” looks like: adherence to a plan, respect for stop-loss placements, emotional regulation after losses, and disciplined trade spacing. Traders who disagree with this model may find the Rating unhelpful, and the system acknowledges this by providing an opt-out mechanism for its most consequential feature (the interface density transition, discussed in Section 6).

3. Rating Computation

3.1 Initialization

Every trader begins with a Rating of 1200. This value is inherited from the Elo tradition, where 1200 represents an average or unproven player. In the DBN context, 1200 carries no claim about the trader’s actual skill level; it simply provides a neutral starting point from which the Rating can move in either direction based on observed behavior and outcomes. The system enforces a floor of zero – ratings cannot become negative – but places no ceiling on the upward range.

3.2 Per-Trade Delta Calculation

When a trade is closed and logged to the centralized data store, the Rating update function computes a delta value that is added to the current Rating. The delta is the sum of an outcome component and up to three behavioral modifier components. The computation proceeds as follows.

The outcome component contributes +5 for a winning trade (defined as a trade with positive realized profit-and-loss) and -3 for a losing trade (defined as a trade with

negative realized profit-and-loss). Trades that close at exactly breakeven contribute zero to the outcome component. This asymmetry – the win bonus exceeding the loss penalty by a ratio of 5:3 – is discussed in detail in Section 4.

The first behavioral modifier addresses revenge trading. The system examines the trade immediately preceding the current trade in the centralized store. If that prior trade was a loss, and the time elapsed between the prior trade’s exit and the current trade’s entry is less than 120,000 milliseconds (two minutes), the current trade is classified as a revenge trade and incurs a penalty of -10. This penalty is applied regardless of the current trade’s outcome. A revenge trade that happens to be profitable still receives the -10 modifier, producing a net delta of -5 (the +5 win bonus minus the -10 revenge penalty). The two-minute threshold is calibrated to the simulation tick rates used in the platform’s trading environments, where one simulated minute maps to one real second at standard speed. In a live trading context, two minutes represents an implausibly short interval for completing a proper pre-trade analysis after absorbing a loss.

The second behavioral modifier addresses stop-loss modification. If the trade object carries a flag indicating that the stop-loss was moved after initial placement, a penalty of -5 is applied. This modifier targets the well-documented behavioral pattern of traders widening their stop-losses to avoid taking planned losses, a behavior that transforms a bounded-risk trade into an unbounded-risk trade and that correlates strongly with eventual account failure. As with the revenge modifier, this penalty is applied regardless of trade outcome.

The third behavioral modifier addresses plan adherence. If the trade carries an adherence score (a percentage reflecting how many items on the pre-trade checklist were completed) and that score is 80% or higher, a bonus of +3 is applied. This is the only behavioral modifier that is positive, reflecting the system’s design as a discipline-reinforcement mechanism rather than a purely punitive one.

The final Rating is computed as $\max(0, \text{previous_rating} + \text{delta})$, where delta is the sum of the outcome component and all applicable behavioral modifiers. The $\max(0, \dots)$ guard prevents negative ratings while allowing the delta itself to be arbitrarily negative.

3.3 History Tracking

Each Rating update appends an entry to a rolling history array stored in the behavioral data structure. Each entry contains a timestamp, the new Rating value, the delta applied, and a human-readable reason string concatenating all applicable modifiers (e.g., “+5 win, -10 revenge”). The history array is capped at 200 entries; when an entry is appended that would exceed this limit, the oldest entry is discarded. This sliding window ensures bounded memory consumption while retaining sufficient history for the regression detection algorithm described in Section 7.

4. Asymmetric Penalty Rationale

The Rating system's penalty structure is deliberately asymmetric along two axes. The first asymmetry is between wins and losses: a winning trade contributes +5 while a losing trade costs only -3. The second, more consequential asymmetry is between outcome penalties and behavioral penalties: the worst possible outcome penalty is -3, while the worst possible behavioral penalty is -10 (revenge trading) or -5 (stop-loss movement).

The win/loss asymmetry of 5:3 serves a specific purpose. In a system where a random 50% win rate is the null hypothesis, the asymmetry produces a slight positive drift for a behaviorally neutral trader. A trader who wins and loses in equal proportion, with no behavioral modifiers triggered, accumulates a net +2 per pair of trades ($+5 - 3 = +2$). This drift is intentional. The Rating is not designed to converge to 1200 for an average trader; it is designed to rise steadily for a trader who avoids behavioral errors, even if their outcome profile is mediocre. The 5:3 ratio ensures that the Rating trends upward for any trader with reasonable discipline, creating a positive reinforcement gradient that rewards consistency.

The more striking asymmetry is between outcome magnitude and behavioral magnitude. A single revenge trade costs -10, which is more than three times the penalty of a simple losing trade (-3) and double the penalty of a moved stop-loss (-5). This encoding reflects an empirical observation about the causal structure of trading failure. A single losing trade is a normal, expected event that carries no information about the trader's long-term trajectory. A revenge trade, by contrast, is a strong signal that the trader has entered a state of emotional dysregulation that typically produces a cascade of further poor decisions. The elevated penalty is not a moral judgment about revenge trading; it is a statistical bet that a revenge trade is a better predictor of near-future negative outcomes than a simple loss is.

The stop-loss movement penalty occupies an intermediate position (-5) between the revenge penalty (-10) and the loss penalty (-3). Moving a stop-loss is a less acute behavioral failure than revenge trading - it does not necessarily indicate emotional dysregulation - but it represents a violation of the trader's own risk management plan that, if habitual, erodes the statistical edge that the plan was designed to capture.

The adherence bonus (+3) is intentionally modest relative to the penalties. The system is designed to penalize bad behavior more aggressively than it rewards good behavior, on the principle that avoiding errors contributes more to long-term trading success than performing any single positive behavior. A trader who simply avoids revenge trading, avoids moving stop-losses, and maintains reasonable plan adherence will see their Rating rise steadily through the win/loss drift alone. The adherence bonus accelerates this rise but is not the primary driver.

The net effect of these asymmetries is that a trader's Rating trajectory is dominated by their behavioral profile rather than their outcome profile. Two traders with identical win rates will diverge sharply in Rating if one engages in revenge trading and the other does not. This is the central design intent: the Rating measures character, not luck.

5. The Density-Rating Bridge

The Trader Rating's most visible consequence is its control over the platform's interface density. The DBN platform defines two interface density modes: "full" and "minimal." Full mode presents all available UI elements - sidebar panels, secondary metrics, behavioral coaching tools, diagnostic analytics, and extended navigation. Minimal mode collapses sidebars, hides secondary metrics, compacts navigation, and presents only the information essential for trade execution and primary performance monitoring.

The transition from full to minimal mode is triggered when a trader's Rating crosses 1600 from below. This threshold was selected to require sustained disciplined trading rather than a brief winning streak. Starting from 1200, a trader accumulating a net +2 per trade pair (one win, one loss, no behavioral flags) requires approximately 200 trades to reach 1600. With behavioral bonuses from consistent adherence, this number decreases, but the order of magnitude remains in the hundreds. The threshold thus represents a meaningful body of evidence that the trader has internalized the behavioral patterns the platform is designed to teach.

The pedagogical reasoning behind the density transition is rooted in the progressive disclosure principle articulated in the platform's core design philosophy. A new trader benefits from the full density of the interface: the behavioral coaching, the diagnostic analytics, the pre-trade gates, and the expanded metric panels provide the scaffolding necessary for developing disciplined habits. An experienced, disciplined trader - as evidenced by a sustained Rating above 1600 - no longer needs this scaffolding. For such a trader, the full-density interface becomes noise rather than signal, adding cognitive load without adding decision-making value. The minimal mode strips the interface to its essential execution surface, trusting the trader to maintain the behavioral patterns they have demonstrated.

The transition is subject to three guards. First, traders operating in Pro-Pass mode (a focused interface designed specifically for prop firm challenge attempts) are exempt from the density transition entirely, as their interface is separately controlled by the Pro-Pass system. Second, traders who have explicitly opted out of the density transition via a settings flag (`minimalOptedOut`) are not transitioned, even if their Rating crosses 1600. This opt-out is surfaced as an "Undo" button in a toast notification that appears at the moment of transition, giving the trader immediate agency to reverse the change. Third, the transition only fires on the exact crossing event - a trader whose Rating is already above 1600 when the system is initialized does not experience a re-transition.

The density transition is not merely cosmetic. Across the four primary platform surfaces - workspace, journal, psychology dashboard, and backtesting - the minimal mode applies targeted CSS rules via a `body.density-minimal` class that hides specific DOM elements using `nth-child` selectors. In the workspace, the journal feed, coaching panel, and killzone insights panel are hidden. In the journal, the stat card grid is reduced from five cards to three. In the psychology dashboard, the status bar is truncated and score ring visualizations are scaled down. The specific elements hidden in each surface were selected based on the assessment that they serve a pedagogical function for developing traders but do not contribute to the decision-making

process of a trader who has already internalized the underlying lessons.

6. Regression Detection and Recovery

The density transition is not a one-way door. The system includes a regression detection mechanism that monitors for sustained behavioral deterioration and automatically reverses the transition to full-density mode when deterioration is detected. This mechanism is the critical safety net that makes the auto-transition defensible: without it, a trader who experiences a behavioral regression after crossing 1600 would lose access to the diagnostic tools they need precisely when they need them most.

Regression detection runs after every Rating update and evaluates three independent triggers. If any trigger fires while the interface is in minimal mode, the system immediately transitions back to full mode and emits a diagnostic event that surfaces a banner to the trader explaining which trigger or triggers activated.

The first trigger monitors revenge trade frequency. The system examines all trades from the past five days (432,000,000 milliseconds) and counts revenge trades using the same two-minute gap heuristic used in the Rating update. If three or more revenge trades are detected in the five-day window, the trigger fires. This threshold reflects the judgment that isolated revenge trades may occur even in disciplined traders during periods of unusual stress, but three or more in a five-day period constitute a pattern that warrants intervention.

The second trigger monitors Rating trajectory. The system examines the Rating history entries from the past seven days and identifies the peak Rating within that window. If the current Rating is more than 75 points below the seven-day peak, the trigger fires. A 75-point drop from a peak above 1600 represents a substantial behavioral regression - recall that the maximum single-trade negative delta is -18 (a losing revenge trade with a moved stop-loss: -3 - 10 - 5), so a 75-point drop requires either multiple severe behavioral violations or a sustained pattern of moderate violations. The seven-day lookback window was selected to capture trends that are more than noise but less than the trader's entire history.

The third trigger monitors absolute Rating level relative to the minimal-mode unlock threshold. If the trader has previously unlocked minimal mode (recorded via a timestamp in the settings) and their current Rating has fallen below 1525, the trigger fires. The 1525 threshold is 75 points below the 1600 unlock threshold, providing a symmetry with the second trigger's 75-point drop criterion. This trigger catches the case where a trader's Rating erodes gradually over weeks rather than dropping sharply in a single week, which might not fire the second trigger due to the seven-day window.

When regression is detected, the system sets the interface density to "full," persists the change, and emits a `density:changed` event with a payload that includes the reason string "regression" and an array of the specific triggers that fired. Consuming pages respond to this event by applying the full-density CSS class, rendering a diagnostic banner that names the specific triggers (e.g., "Regression detected: 4 revenge trades in 5 days; Rating dropped 82 points from weekly peak"), and re-enabling all pedagogical UI elements.

The recovery path is implicit rather than prescriptive. The system does not impose additional requirements for re-unlocking minimal mode beyond the original threshold: the trader simply needs to raise their Rating back above 1600 via sustained disciplined trading. The regression detection mechanism does not lower the unlock threshold or create a higher bar for re-entry. This design choice reflects the principle that the Rating itself is the sufficient evidence of recovery – a trader who has restored their Rating to 1600 has, by definition, demonstrated the behavioral consistency that justifies the minimal interface.

7. Marketplace Integration

The Trader Rating extends beyond the individual trader’s interface to serve as a trust and quality signal in the platform’s Strategy Marketplace. The marketplace allows traders to share custom-built trading strategies with the community, and the Rating system governs who can publish, how published strategies are ranked, and how authors are credentialed.

Publication is gated at a Rating floor of 1800. A trader whose Rating is below 1800 cannot submit strategies to the marketplace; the publish button is visually disabled and displays a tooltip indicating the current Rating and the required threshold. This gate serves a quality-control function: a Rating of 1800 requires not only sustained profitability but sustained behavioral discipline, ensuring that strategies in the marketplace are authored by traders who have demonstrated process quality over a meaningful sample of trades.

Author credentialing uses a two-tier verified badge system. Authors with a Rating of 2000 or above receive a green verified checkmark displayed alongside their name in marketplace listings and leaderboard entries. Authors with a Rating of 2200 or above receive a gold verified checkmark. These tiers provide a visual trust signal to traders browsing the marketplace, allowing them to assess author credibility at a glance without needing to inspect the author’s full trading history.

Strategy ranking in marketplace search results uses a composite score that incorporates the author’s Rating alongside engagement metrics. The ranking formula is:

$$\text{ratingWeightedScore} = \text{downloads} * \text{stars} * \log(\max(\text{authorRating}, 1601) / 1600)$$

where `downloads` is the strategy’s download count, `stars` is its average star rating (1-5), and `authorRating` is the author’s current Trader Rating. The logarithmic transformation of the Rating ensures that Rating contributes meaningfully to ranking without dominating it. At Rating 1800 (the publication floor), the log factor is approximately 0.117. At Rating 2200, it rises to approximately 0.318. At Rating 3000, it reaches approximately 0.631. This scaling means that a very high-rated author receives roughly a 2-3x ranking multiplier compared to a minimally qualified author, which is substantial but does not override a large disparity in download count or star rating. The `max(authorRating, 1601)` guard ensures the logarithm is always positive, preventing division-by-zero and negative scores.

The marketplace also maintains a private library for strategies authored by traders

whose Rating has fallen below the publication floor after initial publication. These strategies are not removed but are moved to a reduced-visibility section with lowered opacity, preserving the author's work while reflecting the reduced trust signal. This mechanism prevents the abrupt disappearance of strategies that other traders may have adopted, while maintaining the integrity of the public marketplace.

8. Volatility Dampening and Stability

Several design choices contribute to the Rating's stability and resistance to noise.

The history cap of 200 entries serves a dual purpose. Its primary function is memory management: in a client-side application persisting state to `localStorage`, unbounded array growth would eventually degrade performance and risk exceeding storage quotas. Its secondary function is implicit recency weighting. Although the Rating itself is a simple running sum with no explicit decay, the regression detection algorithm operates on windows drawn from the history array. By capping the array at 200 entries, the system ensures that only recent trading behavior influences regression detection, even as the Rating itself carries the full cumulative history of all trades.

The floor at zero prevents a negative feedback spiral in which a deeply negative Rating becomes practically irrecoverable. While the floor is unlikely to be reached in normal operation - reaching zero from 1200 requires a net delta of -1200, which at the worst-case per-trade delta of -18 would require approximately 67 consecutive maximally penalized trades - the guard provides mathematical safety against edge cases.

The absence of a K-factor or dynamic scaling is a deliberate simplification. In traditional Elo systems, the K-factor controls the magnitude of rating changes and is often varied based on the player's experience level or the significance of the game. The DBN Rating uses fixed deltas rather than proportional adjustments, which means that a trade at Rating 1200 changes the Rating by the same absolute amount as a trade at Rating 2500. This choice was made for transparency: traders can mentally compute the impact of any given trade without needing to understand a scaling function. The trade-off is reduced discrimination at extreme Rating values, where the fixed deltas represent a smaller proportional change. This trade-off is acceptable because the Rating's primary function is to drive the density transition and marketplace gates, both of which operate at fixed thresholds rather than relative positions.

The system does not implement matchmaking or opponent-relative adjustments because there is no opponent. Each trader is evaluated against an absolute behavioral standard rather than relative to other traders. This distinguishes the system from true Elo implementations and is the primary reason the system is described as "ELO-inspired" rather than "ELO-based." The absence of opponent modeling eliminates the need for provisional ratings, K-factor schedules, or deflation corrections that Elo implementations in competitive contexts must address.

9. Validation Methodology

Rigorous validation of the Rating system requires establishing two empirical claims: first, that the Rating correlates with long-term trading outcomes more strongly than outcome-only metrics; second, that the regression detection mechanism accurately identifies traders whose future performance will deteriorate.

The first claim can be tested via a longitudinal cohort study. Traders are segmented at a fixed point in time by their Rating and by their trailing 30-day win rate. For each segmentation, the subsequent 90-day equity trajectory is measured. The hypothesis is that Rating-based segmentation produces tighter within-group variance and wider between-group separation than win-rate-based segmentation. A Rating of 1600+ should predict positive 90-day equity change more reliably than a win rate of 55%+, because the Rating incorporates behavioral consistency information that the win rate does not.

The second claim can be tested by measuring the regression detection mechanism's sensitivity and specificity. True positives are traders who experienced a regression trigger and subsequently suffered a drawdown exceeding 10% of peak equity within 30 days. False positives are traders who triggered but did not suffer the drawdown. The target sensitivity is above 70% with a specificity above 60%.

As of this writing, the platform has not yet accumulated sufficient longitudinal data to conduct either study at statistical power. The Rating system is deployed and collecting data, and the validation studies are planned for execution once the platform's active user base reaches a scale sufficient for meaningful cohort analysis. The thresholds and penalty magnitudes described in this paper should therefore be understood as initial calibrations subject to revision based on empirical findings.

10. Limitations and Open Questions

The Rating system has several acknowledged limitations that merit discussion.

The cold-start problem is the most immediate. A new trader begins at 1200 with no history, and the system cannot distinguish between a genuinely new trader and an experienced trader who has created a new account. The first 50-100 trades represent a calibration period during which the Rating is noisy and unreliable. The platform partially mitigates this through a separate calibration quiz system that initializes other subsystems, but the Rating itself begins from the same 1200 baseline for all users. A future enhancement could allow the calibration quiz to adjust the initial Rating based on demonstrated knowledge, though this introduces the risk of overfitting to quiz performance rather than live trading behavior.

Small sample bias affects the behavioral modifiers disproportionately. The revenge trade detection relies on sequential trade pairs, meaning a single coincidence - two trades close in time for legitimate reasons - can trigger a -10 penalty. Over many trades, such coincidences wash out. Over few trades, they can dominate the Rating trajectory.

Gaming potential exists in theory. A trader who understands the formula could inflate their Rating by spacing trades widely, never moving stop-losses, and checking adherence checklists mechanically. The system's defense is that the behavioral patterns it incentivizes are genuinely beneficial regardless of motivation. A trader who "games" the Rating by consistently following these patterns is, in practice, trading well.

The fixed delta structure, while transparent, lacks the adaptive properties of true Elo systems. In competitive Elo, upsets produce larger rating changes than expected results, accelerating convergence to true skill. The DBN Rating has no such mechanism: a +5 win is a +5 win regardless of trade quality. A future iteration could scale the win bonus by the achieved risk-to-reward ratio, but this would add complexity to a system whose transparency is a feature.

The relationship between the Rating and the Lock predictive gating system (a separate subsystem not detailed in this paper) creates potential interaction effects. Traders who filter entries through the Lock will take higher-quality setups by construction, producing better outcomes and higher Ratings. This positive feedback loop is generally desirable but could mask situations where the Lock's weights are miscalibrated - the trader's Rating rises because of the Lock's filtering, not because they have independently developed the judgment the Rating is intended to measure.

The interface density transition is a binary switch rather than a continuous spectrum. A trader at Rating 1599 sees the full interface; a trader at 1601 sees the minimal interface. This step function creates a potential oscillation problem for traders whose Ratings hover near 1600, though the 75-point hysteresis between the unlock threshold (1600) and the regression re-lock threshold (1525) mitigates this in practice.

11. Invitation for Peer Review

The DBN Trader Rating system represents an initial attempt to formalize the long-standing intuition in trading psychology that process quality and outcome quality, while correlated in the long run, carry independent information that should be tracked independently. The system is deployed, collecting data, and influencing the user experience of an active trading education platform. Its thresholds, penalty magnitudes, and behavioral detection heuristics are initial calibrations informed by domain expertise and internal testing rather than large-scale empirical optimization.

The authors invite review and critique from practitioners in quantitative finance, behavioral economics, human-computer interaction, and trading psychology. Specific areas where external input would be valuable include: the asymmetric penalty magnitudes, which were set by judgment rather than optimization and would benefit from a principled empirical calibration framework; the revenge trade detection heuristic, which proxies emotional dysregulation via timing and could potentially be improved through input latency analysis; and the regression detection triggers, whose false-positive rate in production has not yet been measured.

The Rating system's source code is available for inspection in the DBN platform's centralized data engine (`dbn-core.js`), where the `updateRating()` and `checkRegression()` methods implement the algorithms described in this paper. The authors wel-

come correspondence directed to the technical contact listed in the platform's project documentation.

*Drive By Numbers (DBN) - drivebynumbers.com Technical White Paper Series, No. 1
April 2026*