Human Day Trading Versus Fly-Derived Reservoir Trading:

A Deep Comparative Study of Returns, Return Structure, Risk Appetite, and a Prospective Single-Trader Test

Ogdn Ames

Arizona State University

Ames Investment Systems

September 14, 2026

Author Note

Ogdn Ames is the author of this manuscript. The paper synthesizes two prior manuscripts by the author: a MaleCNS-derived temporal-memory and blackjack study and a retrospective fly-derived momentum-trading study. The current manuscript was prepared with AI-assisted literature retrieval, drafting, quantitative cross-checking, figure generation, and APA formatting. No peer review or journal acceptance is claimed.

A critical evidentiary boundary applies to the human comparison. No transaction ledger, brokerage statement, order log, or synchronized market-data record for a new individual human day trader was supplied for this manuscript. Accordingly, no individual human P&L, Sharpe ratio, return distribution, or behavioral coefficient is invented. The human side of the present study is a literature-grounded empirical benchmark and a fully specified prospective single-trader protocol. A true head-to-head performance claim requires a future experiment in which the human and fly-derived systems face the same market stream, timing rules, risk limits, and execution model.

A second boundary is temporal. The prior fly-derived financial study used daily total-return-index data; it did not execute an intraday day-trading test. Therefore, the present paper compares economic evidence and return/risk structure while treating direct fly-versus-human performance ranking as unresolved. The preferred next experiment is an intraday rerun of the fly-derived model synchronized to the human trader.

Abstract

This paper compares a fly-connectome-derived artificial recurrent trading system with the empirical behavior of human day traders, focusing on returns, return structure, and risk appetite. The fly evidence comes from two prior manuscripts: a replicated MaleCNS-derived reservoir study showing that anatomical topology can matter under a fixed artificial regime while dynamical timescale and readout capacity dominate absolute counting error, and a retrospective daily momentum-trading study in which the fly family did not establish a benchmark-beating trading advantage. In the common daily period, the fly family trailed buy-and-hold on the assigned mean fold objective for SPY, GLD, and QQQ; AAPL's small mean-of-ratios advantage reversed under equal-capital aggregation. Human evidence is also unfavorable on average but substantially more heterogeneous. U.S., Taiwanese, and Brazilian studies generally find that most day traders lose after costs, while a small right tail demonstrates persistent skill. Unlike the fly-derived policy, however, human risk taking is endogenous and path dependent: prior gains and losses, overconfidence, sensation seeking, affect, and familiarity can change subsequent exposure and turnover. The central conclusion is therefore not that either humans or fly-derived computation 'wins.' The two systems currently answer different questions. The fly model implements an engineered state-to-risk mapping; humans exhibit adaptive, psychologically mediated risk choice. A valid comparison requires a prospective intraday experiment with matched data, constraints, transaction costs, and preregistered outcomes. A 252-session single-human protocol is specified for that purpose.

Keywords: day trading, reservoir computing, Drosophila connectome, behavioral finance, risk appetite, return structure, transaction costs, prospective validation

Human Day Trading Versus Fly-Derived Reservoir Trading

A useful comparison between a human day trader and a fly-derived artificial network has to begin by rejecting the most seductive version of the question. A connectome-derived reservoir is not a fly that experiences greed, fear, loss aversion, confidence, fatigue, or a desire to break even. It is an engineered recurrent state machine whose topology is borrowed from anatomy and whose market actions are produced by imposed dynamics, features, readouts, thresholds, and risk rules (Ames, 2026a, 2026b). A human trader, by contrast, is a biological decision maker whose beliefs, attention, affect, prior outcomes, and risk preferences can change during the trading process. Consequently, the scientifically valid comparison is between realized decision structures under matched constraints, not between human psychology and an imagined 'fly psychology.'

This distinction matters because both prior manuscripts already show that attractive narratives can outrun the evidence. In the counting study, the native MaleCNS topology was useful under a prospectively fixed normalized regime, but the dominant improvement in absolute accuracy came from longer artificial memory and a larger readout applied to all graph families; the exact scalar accumulator remained perfect (Ames, 2026b). In the trading study, the fly-derived architecture produced a reproducible computational pipeline, but the economic result was negative: it did not establish that anatomical topology improved risk-adjusted returns relative to buy-and-hold or calibrated controls (Ames, 2026a). The appropriate new question is therefore narrower: how does the fly-derived policy's observed return/risk structure compare with what is known about human day-trading return distributions and risk taking, and what experiment would allow an actual new human trader to be tested fairly against it?

Research Questions and Evidentiary Standard

The study addresses four questions. First, what did the prior fly-derived experiments actually establish about memory, allocation, returns, and risk? Second, what does primary empirical research establish about human day-trader profitability and the cross-sectional distribution of skill? Third, how does human risk appetite differ structurally from the fly-derived system's engineered exposure policy? Fourth, what prospective design would permit a genuine single-human-versus-fly comparison without fabricating outcomes or exploiting historical hindsight?

The evidentiary standard is deliberately asymmetric. Claims about the fly system are drawn directly from the two supplied manuscripts. Claims about humans are drawn from primary or near-primary academic sources, with emphasis on transaction-level studies of U.S., Taiwanese, Brazilian, futures, and foreign-exchange traders. The literature search is a targeted evidence synthesis, not a formal systematic review or meta-analysis. No pooled human success rate is calculated because the underlying studies use different markets, time periods, trader definitions, cost conventions, and outcome thresholds. Where numerical comparisons are shown together, they are explicitly labeled as noncommensurate descriptive benchmarks.

Table 1

Evidentiary Boundary for the Comparison

Dimension	Prior fly trading study	Human day-trader literature	Prospective matched test
Trading horizon	Daily; close-to-close total-return-index implementation	Intraday; minutes to hours; transaction-level in many studies	Intraday; same bars/quotes and session boundaries for both systems
Observed subject	Artificial reservoir using MaleCNS-derived topology	Many human traders across multiple markets	One new human trader plus frozen fly-derived systems
Returns	Retrospective simulated daily portfolio returns	Realized or reconstructed trader P&L, often net of observed/estimated costs	Synchronized net daily account returns under one execution model
Risk	Exposure, turnover, drawdown, cost scenarios	Leverage/size, turnover, loss chasing, house-money effects, emotion, attention	Gross/net exposure, turnover, tail loss, post-gain/post-loss risk changes
Prospective status	Frozen future protocol prepared; no untouched market outcome in supplied study	Mostly historical observational evidence	Required before any head-to-head performance claim
Direct “winner” claim	Not supported	Not applicable	Permitted only after matched prospective data

Note. The existing fly implementation and human day-trading studies operate at different horizons. Direct performance ranking is therefore invalid until the market stream and rules are matched.

What the Fly-Derived Research Established

Connectome-Derived Memory Is Real but Artificial

The MaleCNS study used a fixed 5,000-neuron induced subgraph from the released male Drosophila central nervous system connectome and imposed artificial recurrent dynamics on that graph. In the independent follow-up, ten new computational seeds, 2,048 training shoes, 512 validation shoes, and 1,024 locked test shoes were used. Under the fixed postsynaptically normalized configuration, native models achieved a mean pooled true-count RMSE of 0.8085 versus 0.9990 for uniform random models and 1.1572 for the degree/raw-strength null. Native topology won the fixed native-versus-random contrast in 9 of 10 computational seeds (Ames, 2026b).

That topology result did not imply that anatomy was the main source of computational accuracy. A validation-only engineering extension lengthened the artificial memory timescale and expanded the sampled readout to 256 states. The selected native family reached a mean true-count RMSE of 0.0097, selected random networks reached 0.0150, and the selected raw-strength null reached 0.0190. All three selected families achieved 100% nearest-integer running-count accuracy on the finite fresh-shoe sample, while the exact one-state accumulator remained error-free. Thus, native topology affected performance within a fixed regime, but model dynamics and readout capacity produced the dominant absolute improvement (Ames, 2026b).

This distinction is directly relevant to finance. If a fly-derived trader outperforms a control, one still has to identify whether the increment comes from biological graph structure, effective signed strengths, memory timescale, readout capacity, feature transformations, or a decision gate. The counting study shows why topology cannot be treated as a magic label attached to an otherwise conventional signal engine.

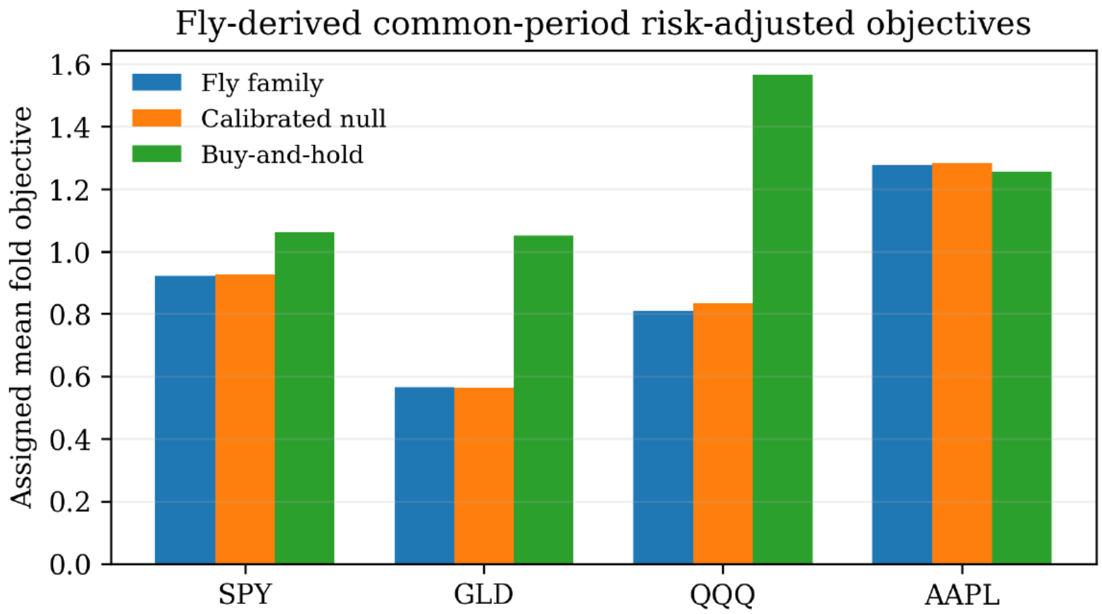
The Existing Financial Test Was Daily, Retrospective, and Economically Unfavorable

The fly-derived trading paper built a nonlinear momentum-and-risk allocation system around the same general reservoir concept. It used 5-, 20-, and 60-bar momentum features scaled by recent volatility, a leaky recurrent update, a seeded readout gate, and a policy bank producing clipped market exposures. The market source was a licensed Bloomberg total-return-index cache. SPY, QQQ, and AAPL covered September 2023 through September 2026, GLD had a longer history, and the common comparison window ran from September 2023 through September 2026. Intraday cells were not available; the executed financial evidence was daily only (Ames, 2026a).

The paper's primary economic result was negative but informative. On the common period, arithmetic mean fold objectives for fly / calibrated null / buy-and-hold were 0.9212 / 0.9260 / 1.0630 for SPY Sharpe; 0.5654 / 0.5624 / 1.0521 for GLD Sharpe; 0.8100 / 0.8337 / 1.5656 for QQQ Sortino; and 1.2769 / 1.2844 / 1.2564 for AAPL Sortino. The small AAPL mean-of-ratios advantage disappeared when returns were combined in an equal-capital ensemble. No multiplicity-adjusted comparison established superiority (Ames, 2026a).

Figure 1

Fly-Derived Common-Period Risk-Adjusted Objectives



SPY/GLD use Sharpe; QQQ/AAPL use Sortino. Values are mean fold ratios, not pooled returns.

Note. Source values are from Ames (2026a). SPY and GLD use Sharpe ratios; QQQ and AAPL use Sortino ratios. These are arithmetic means of fold ratios and should not be interpreted as pooled continuous-account performance.

Table 2

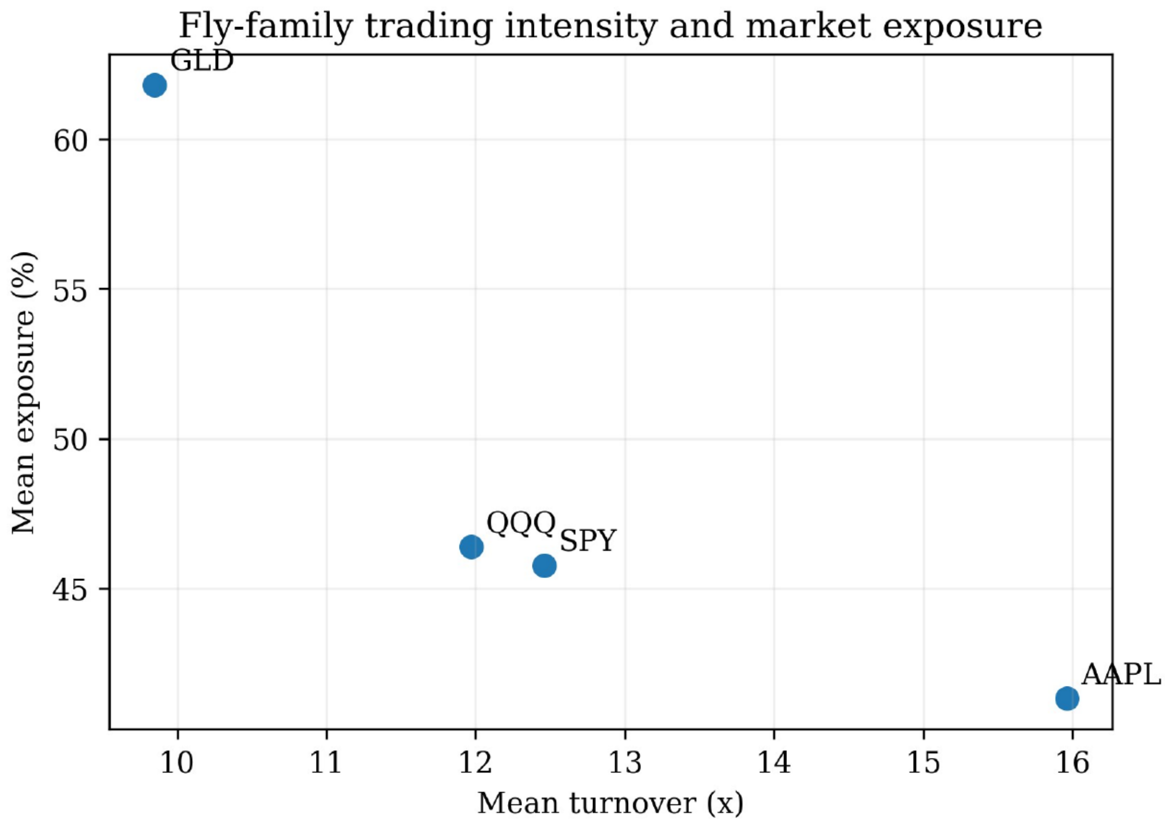
Fly-Family Common-Period Trading and Risk Summary

Asset	Assigned objective	Fly	Buy-and-hold	Fly - B&H	Profit / $100k	Max drawdown	Exposure	Turnover
SPY	Sharpe	0.9212	1.0630	-0.1418	$3,976	-3.84%	45.77%	12.46x
GLD	Sharpe	0.5654	1.0521	-0.4867	$13,468	-10.32%	61.81%	9.85x
QQQ	Sortino	0.8100	1.5656	-0.7556	$2,441	-6.89%	46.39%	11.97x
AAPL	Sortino	1.2769	1.2564	+0.0205	$6,299	-9.15%	41.33%	15.96x

Note. Source: Ames (2026a). Profit, drawdown, exposure, and turnover are fold/recurrent-family summaries, not a continuously compounded live brokerage account. The AAPL mean-of-ratios advantage reverses under the paper's equal-capital ensemble analysis.

Figure 2

Fly-Family Trading Intensity and Market Exposure



Note. Source values are from Ames (2026a). The four observations are descriptive asset-level summaries. Turnover and exposure are economically important because a strategy can obtain a lower drawdown simply by holding more cash; lower drawdown alone does not imply superior timing.

Relative to buy-and-hold, the fly assigned objective was approximately 13.3% lower for SPY, 46.3% lower for GLD, and 48.3% lower for QQQ; AAPL was approximately 1.6% higher on the non-pooled mean-of-ratios statistic. Those percentage differences are descriptive transformations of heterogeneous Sharpe and Sortino ratios, not a cross-asset performance index. The average fly-family exposure across the four assets was 48.8%, mean turnover was 12.56x, and the average absolute maximum-drawdown statistic was 7.55%. These simple averages help characterize the policy but should not be elevated to inferential endpoints.

Human Day-Trading Returns: Average Losses, Cost Drag, and a Thin Right Tail

Most Samples Show Negative Average Economics

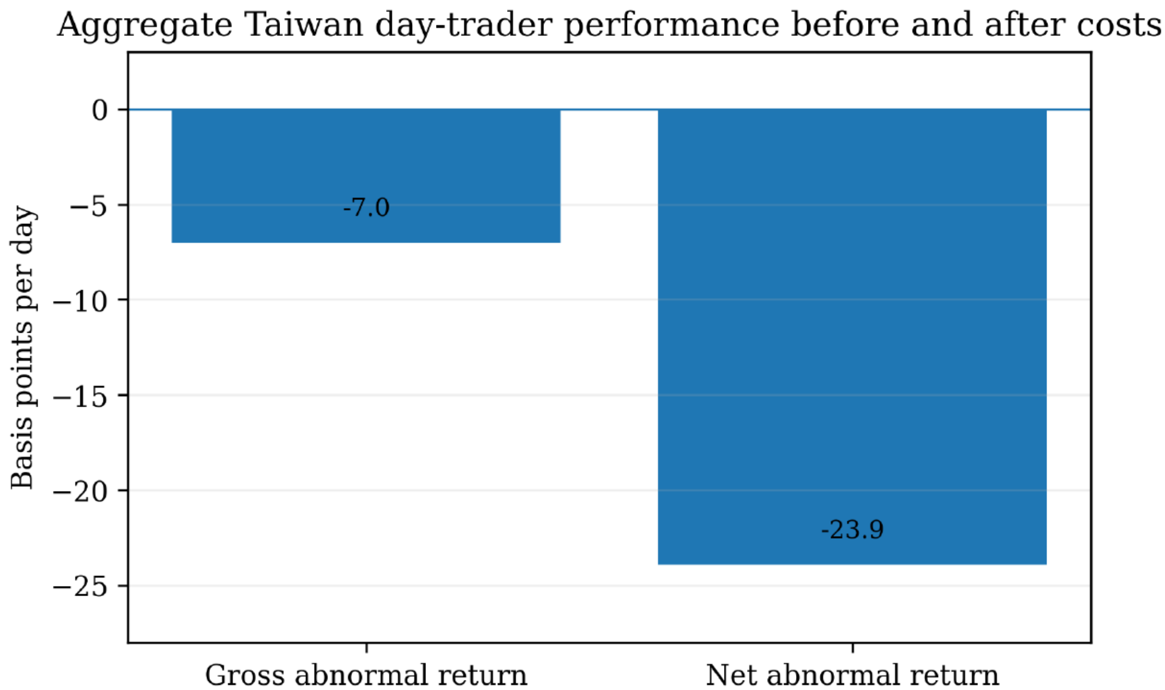
The human evidence is unusually consistent on one point: frequent speculative trading is difficult to monetize after costs. Jordan and Diltz (2003) analyzed 324 U.S. day traders. They found 35.8% with positive net profit and 64.2% with negative net profit after commissions; the average gross profit was above $8,000, while average net profit was approximately -$750. Transaction costs changed the economic conclusion for the average trader. The authors also found day-trader profitability related to Nasdaq movements, underscoring that realized P&L is conditional on market regime rather than an invariant personal trait.

The Taiwan evidence is much larger. Barber, Lee, Liu, and Odean (2014) studied all day traders on the Taiwan Stock Exchange from 1992 to 2006. Roughly 450,000 individuals day traded in an average year, and about 277,000 exceeded NT$600,000 of day-trading volume. Approximately 20% of this active group earned positive abnormal returns net of fees in a given year, but less than 1% of the total day-trader population was able to earn positive abnormal returns predictably and reliably net of fees. When traders were ranked on prior-year performance, the top 500 earned 61.3 basis points per day before fees and 37.9 basis points after fees in the next year, whereas bottom-ranked traders earned -11.5 and -28.9 basis points, respectively (Barber et al., 2014). The return distribution therefore contains genuine heterogeneity: 'most lose' and 'some skill exists' are simultaneously true statements.

The later Taiwan learning study strengthens the negative aggregate conclusion. Barber, Lee, Liu, Odean, and Zhang (2020) report an aggregate abnormal intraday loss of 7 basis points per day before costs and 23.9 basis points per day after costs. Seventy-four percent of day-trading volume was generated by traders with a history of losses, and 97% of day traders were estimated likely to lose money in future day trading. Poor performers were more likely to quit, but many unprofitable traders persisted; the pattern was not consistent with a simple rational-learning story in which experience efficiently removes bad traders.

Figure 4

Aggregate Taiwan Day-Trader Performance Before and After Costs



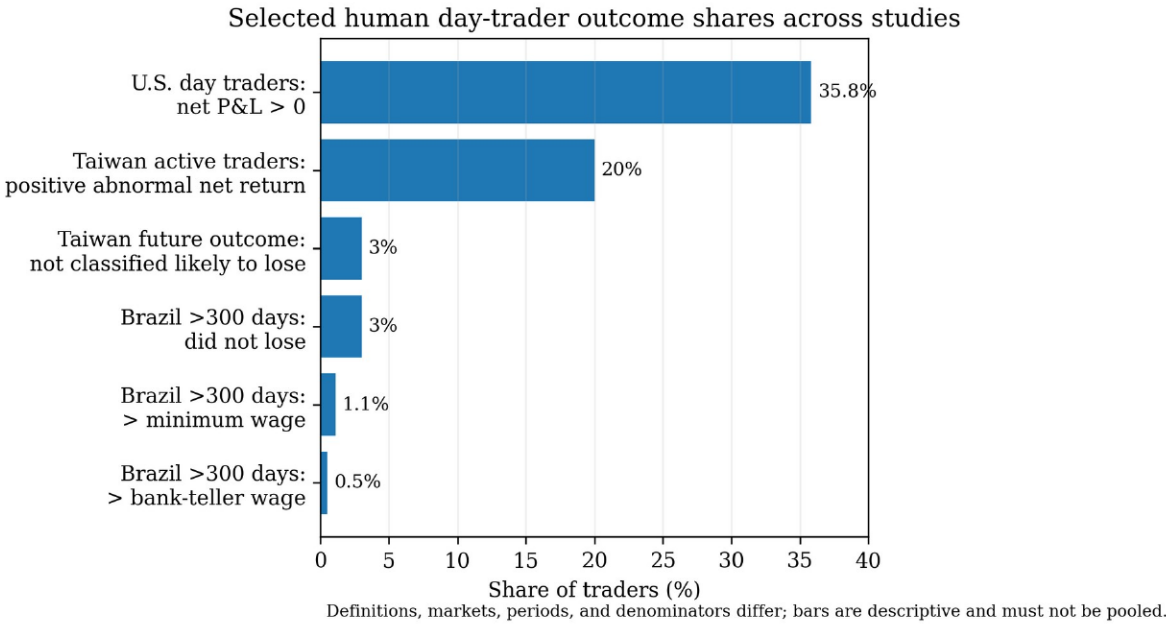
Note. Barber et al. (2020), Taiwan Stock Exchange, 1992-2006.

Note. Values are aggregate abnormal intraday returns from Barber et al. (2020). The 16.9-basis-point gross-to-net wedge illustrates why turnover and execution costs must be part of any human-versus-model comparison.

Brazilian evidence reaches a similar conclusion in a different market structure. Chague, De-Losso, and Giovannetti (2020) examined individuals who began day trading equity futures between 2013 and 2015 and persisted for at least 300 days. Ninety-seven percent lost money; 1.1% earned more than the Brazilian minimum wage, 0.5% earned more than the initial salary of a bank teller, and the top trader earned approximately US$310 per day with a standard deviation of roughly US$2,560. A subsequent study of the COVID-era expansion of Brazilian day trading documents approximately 100,000 active participants per day and R$9.9 billion of gross losses from the pandemic onset through December 2023, about R$10,200 per individual (Chague & Giovannetti, 2025).

Figure 3

Selected Human Day-Trader Outcome Shares Across Studies



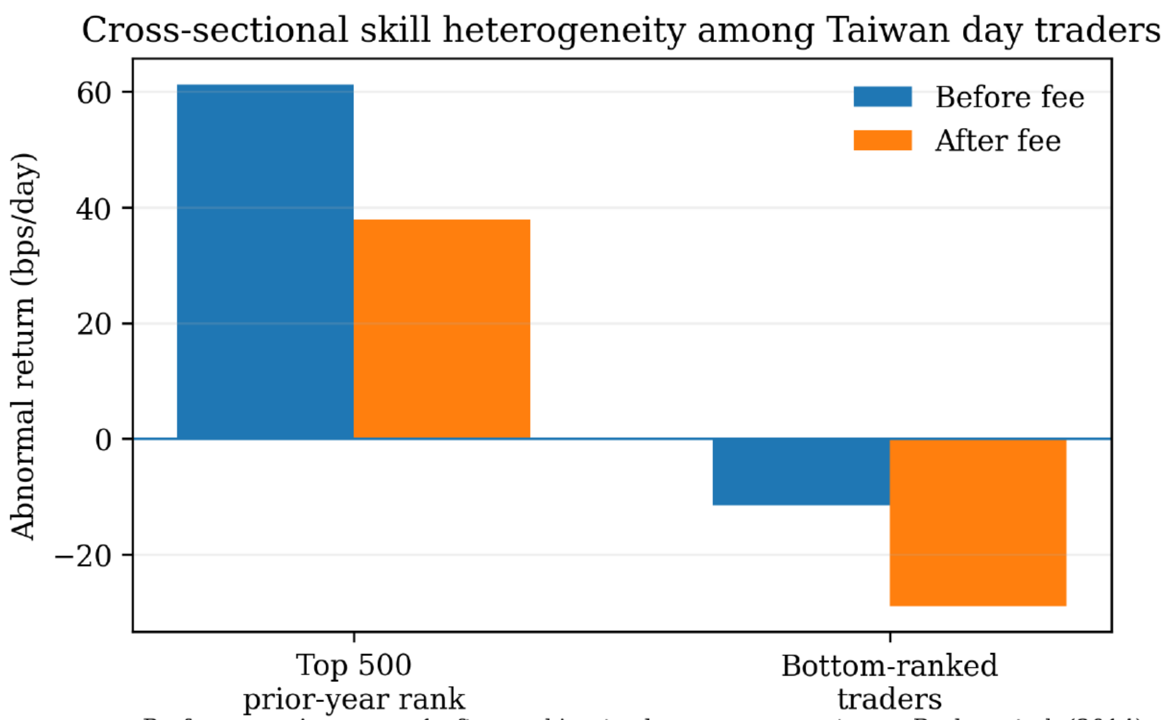
Note. The bars are deliberately not pooled. Jordan and Diltz (2003), Barber et al. (2014, 2020), and Chague et al. (2020) use different markets, samples, cost assumptions, and thresholds. The figure shows how success becomes rarer as the criterion moves from merely positive realized P&L to persistent, economically meaningful profitability.

Persistent Skill Exists, but It Is Sparse

A weak summary of the literature is 'day traders lose.' A stronger summary is distributional. The modal or average experience is unfavorable, transaction costs are economically material, and trader ability is highly dispersed. Barber et al. (2014) show that past performance contains information about future performance at the extreme upper tail. Less than 1% of the day-trader population was reliably profitable after fees, but those traders produced economically large abnormal returns. This is important for the proposed single-human study: population base rates are relevant priors, but they do not determine the performance of a specific preidentified trader. The correct test is prospective, not an inference from population averages to an individual.

Figure 5

Cross-Sectional Skill Heterogeneity Among Taiwan Day Traders



Note. Barber et al. (2014) ranked traders by prior-year performance and measured next-year returns. The result demonstrates persistence in the extreme upper tail while showing severe underperformance in the lower tail.

Table 3

Selected Empirical Evidence on Human Day Trading

Study	Market / sample	Primary evidence	Implication for comparison
Jordan & Diltz (2003)	U.S. equity; 324 day traders	35.8% net positive; 64.2% net negative; mean gross profit > $8,000 but mean net profit about -$750	Execution costs can reverse the average economic conclusion.
Barber et al. (2014)	Taiwan equities; 1992-2006	~20% of active traders positive abnormal net in a year; <1% predictably positive; top 500 +37.9 bps/day net next year	Skill is highly heterogeneous and concentrated in a thin right tail.
Barber et al. (2020)	Taiwan equities; 1992-2006	-7 bps/day gross and -23.9 bps/day net aggregate; 97% likely to lose in future; 74% volume from prior losers	Poor performance persists despite experience; costs are large.
Chague et al. (2020)	Brazil equity futures; new entrants who persisted >300 days	97% lost; 1.1% > minimum wage; 0.5% > bank-teller wage	Persistence alone does not imply learning or profitability.
Chague & Giovannetti (2025)	Brazil exchange; COVID era through 2023	~R$9.9 billion gross losses; ~R$10,200 average loss per individual	Negative economics scale with mass retail participation.
Coval & Shumway (2005)	CBOT proprietary traders	Morning losers ~16% more likely to take above-average afternoon risk	Human risk exposure responds to own prior outcomes.
Lo et al. (2005)	80 day traders over five weeks	More intense emotional reaction to gains/losses associated with worse performance	Affect is a measurable state variable in human trading.
Ben-David et al. (2018)	Retail FX traders	Winning feedback increased trade size, size variability, and trade count even when past performance did not predict future success	Self-attribution can dynamically amplify risk.

Note. The studies are not a common estimand and should not be meta-analytically pooled without harmonized definitions. Their value here is structural: they identify cost drag, performance heterogeneity, persistence, and state-dependent risk behavior.

Return Structure: The Most Important Difference Is Feedback

Fly Returns Are the Output of a Frozen Mapping

The fly-derived system's return process is generated by market inputs passed through an imposed recurrent transform and a fixed decision policy. In the supplied financial study, momentum features over several horizons and recent volatility affected reservoir state; the readout then selected among discrete policy experts and exposure actions (Ames, 2026a). Historical P&L itself was not a psychological input. A loss does not make the model angry, desperate, ashamed, euphoric, overconfident, or determined to get back to even. Unless prior P&L is explicitly supplied as a feature or used to alter a rule, the strategy's next action is not conditioned on the subjective experience of its last outcome.

This has two consequences. First, the fly-derived return distribution can still exhibit serial dependence, volatility clustering, nonlinear exposure, drawdown, and tail risk because the market inputs are serially structured and the recurrent state has memory. Second, these statistical features should not be confused with endogenous risk preference. The recurrent state is memory of the engineered input stream, not a utility function that changes after gains or losses.

Human Returns Are Jointly Market- and Trader-State Dependent

Human return structure contains an additional feedback channel. Coval and Shumway (2005) found that proprietary traders with morning losses were about 16% more likely than morning winners to assume above-average risk in the afternoon, consistent with break-even behavior. Huang and Chan (2014), using Taiwan futures records, found nonlinear effects of morning outcomes on afternoon risk: active individuals displayed a house-money effect after large gains and break-even behavior after losses. Ben-David, Birru, and Prokopenya (2018) found that retail foreign-exchange traders increased trade sizes, trade-size variability, and trading frequency following gains even though past success did not predict future success; the effect was stronger among novices.

These findings make the human P&L process endogenous. Market returns affect trader wealth and emotion; those outcomes change subsequent position size or trade frequency; the changed positions then alter future P&L. A return series is therefore not merely a passive score of signal quality. It is partly the product of a feedback controller whose parameters can shift with recent experience.

Cross-Sectional Preferences and Attention Also Shape the Opportunity Set

Risk appetite is not only a within-day response to P&L. Dorn and Huberman (2005) linked self-reported risk tolerance to less diversified portfolios and more aggressive trading. Grinblatt and Keloharju (2009) found that overconfidence and sensation seeking were associated with higher trading frequency. Odean (1998) documented the disposition effect: investors preferentially realize winners and retain losers, a behavior not explained by rebalancing or subsequent performance. Kumar (2009) found that individual investors disproportionately held lottery-like stocks and that those stocks underperformed, linking speculative security choice to gambling propensity. More recently, Chague, Giovannetti, and Paiva (2026) found that a local physical store more than doubled the likelihood that individuals in small Brazilian cities day traded the firm's stock, consistent with top-of-mind familiarity rather than an intraday information advantage.

The implication for a human-versus-fly experiment is operational: if the human may choose symbols using attention, familiarity, or discretionary beliefs while the model is forced to trade a fixed universe, then differences in returns combine security selection with execution skill and risk taking. A fair test should either freeze the same tradable universe for both participants or explicitly treat security selection as part of the human strategy and give the model the same universe and data access.

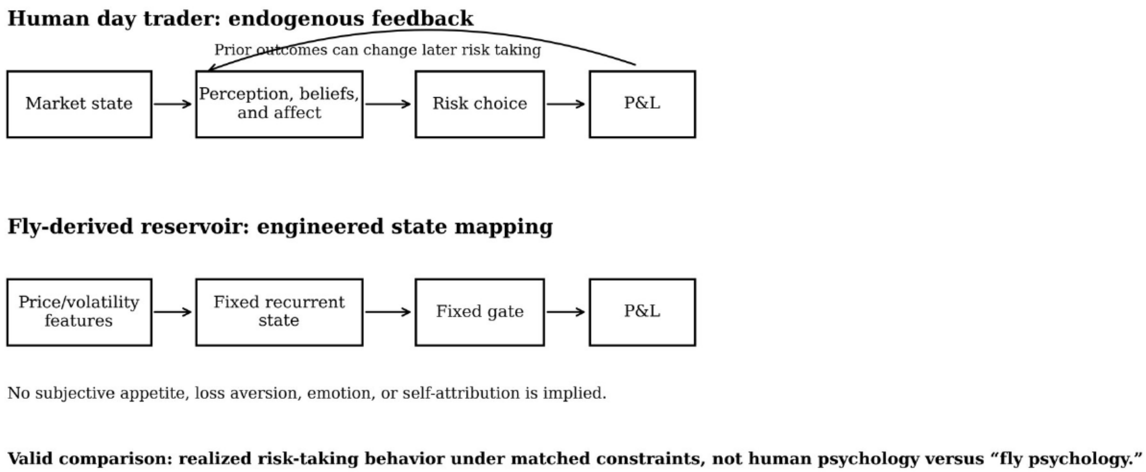
Risk Appetite: Human Construct Versus Model Parameter

The phrase 'risk appetite of the fly' is convenient but technically inaccurate. Appetite implies preference. The artificial reservoir has no subjective preference; it has a risk policy. In the daily fly study, cash assumptions, cost multipliers, gating thresholds, action sizes, and exposure rules were external engineering choices. The model can be more or less risky in observed behavior, but it cannot be said to prefer risk unless preference is defined purely operationally as a mapping from state to exposure.

Human risk appetite is meaningfully different because it can be measured both behaviorally and, imperfectly, psychologically. Dorn and Huberman (2005) show that self-reported risk tolerance predicts turnover and diversification. Lo and Repin (2002) measured physiological responses during live professional trading and found electrodermal responses associated with transient market events and cardiovascular changes associated with market volatility. Lo, Repin, and Steenbarger (2005) then found in 80 day traders that stronger emotional reactions to gains and losses were associated with worse trading performance, while no single personality profile defined a successful trader. Laboratory evidence also indicates that anticipatory neural activity can precede risk-seeking and risk-avoidance errors (Kuhnen & Knutson, 2005). These findings do not turn trading into a neurobiological destiny; they show that human risk choice has latent internal state variables absent from the fly-derived policy as currently implemented.

Figure 6

Human Outcome Feedback Versus Fly-Derived Engineered State Mapping



Note. Conceptual comparison. The fly-derived network can have recurrent memory of price/volatility inputs, but no subjective affect or self-attribution is implied. To test equivalent economic behavior, both systems should be compared under the same risk and execution constraints.

Table 4

Risk-Appetite Signatures and Testable Proxies

Construct	Human trader	Fly-derived system	Prospective measurement
Response to prior loss	Break-even/loss-chasing can increase later risk	No effect unless P&L is an explicit feature or risk rule	Change in gross exposure and trade size after negative prior P&L
Response to prior gain	House-money and self-attribution can increase risk	No subjective gain response	Exposure elasticity after positive prior P&L
Overconfidence	Can increase trading frequency and turnover	Not applicable as a mental state	Calibration of forecast confidence vs outcomes; turnover conditional on confidence
Sensation / entertainment	Can increase trading participation	Not applicable	Optional questionnaire; trade frequency beyond signal opportunities
Loss realization	Disposition effect can delay loss realization	Exit rule is algorithmic	PGR-PLR / holding-time asymmetry for winners vs losers
Attention / familiarity	Can shape symbol choice	Universe is designer-specified	Fraction of discretionary trades in familiar / salient symbols
Physiological / affective state	May covary with risk processing and performance	No biological affect in the artificial model	Optional self-report or wearable data; exploratory only
Risk budget	Can drift intraday without controls	Explicit hard-coded cap	Matched gross leverage, per-symbol caps, no-overnight rule

Note. Human constructs should not be anthropomorphically assigned to the reservoir. The comparison uses observable risk-taking proxies.

A Quantitative Comparison Without a False Head-to-Head Claim

At first glance, the human and fly evidence look similar because both contain unfavorable average economic outcomes. That superficial similarity should not be overstated. The fly study's assigned risk-adjusted objectives compare a daily simulated strategy with buy-and-hold and matched computational controls. The human literature usually compares realized intraday profits or abnormal returns across traders, with different leverage, markets, trading opportunities, and costs. Sharpe, Sortino, abnormal-return alpha, and raw profit are not interchangeable estimands. Nor is a 2023-2026 ETF/stock backtest comparable to a 1992-2006 Taiwan day-trading panel.

What can be compared is structure. First, costs matter in both systems: the fly paper found sensitivity to cost multipliers, while human studies show gross-to-net deterioration large enough to reverse profitability. Second, heterogeneity matters: the fly paper's apparent AAPL success depended on aggregation, while human studies show a small skilled upper tail embedded in a large losing population. Third, evaluation discipline matters: the fly paper's strongest remaining test was prospective; the human literature likewise shows why ex post selection of an exceptional trader is not evidence that a new trader drawn today will be profitable. Fourth, risk should be decomposed into exposure, turnover, drawdown, tail loss, and state dependence rather than summarized by a single ratio.

The sharpest difference is control. The fly policy is comparatively stable because its mapping is frozen. Human behavior can shift after the very outcomes being evaluated. This means a prospective human protocol needs stronger behavioral accounting than a model backtest. Every order, cancellation, size change, leverage change, stop change, and rule exception is part of the strategy and must be recorded rather than normalized away after the fact.

Prospective Single-Human Versus Fly Experiment

Why the Fly Must Be Rerun Intraday

The existing fly evidence cannot be used as the machine side of a fair day-trading contest because it is daily. The clean design is to rerun the fly-derived reservoir on intraday data and synchronize it to the human's opportunity set. The alternative—forcing the human to make only one daily decision—would no longer test day trading. Therefore the proposed study is a new experiment, not a re-labeling of the prior backtest.

Core Design

The recommended protocol covers 252 market sessions, approximately one trading year, with locked descriptive checkpoints after 63 and 126 sessions and no strategy modification at those checkpoints. The tradable universe is SPY, QQQ, GLD, and AAPL to preserve continuity with the prior fly study. Both participants begin with the same nominal equity and are subject to the same hard limits: no overnight positions, identical maximum gross exposure, identical per-symbol exposure caps, no options, no external capital flows, and the same short-sale feasibility rules. If leverage is allowed, the exact leverage cap is fixed in advance and applied symmetrically.

The data stream should be exchange-grade or consolidated intraday trades and quotes, not a total-return index. At minimum, timestamps, bid, ask, last price, and volume should be retained. Corporate actions are handled outside the intraday session. The fly-derived model should operate on a prespecified bar interval such as five minutes, and its features should be reconstructed only from information available at each decision time. The human may trade at any time during the regular session, but performance is reconciled to the same market-data clock. A secondary analysis can discretize human inventory at fly decision times for state-to-state comparisons.

Execution should be deliberately conservative. For a simulated fly order, the fill price should be tied to a future executable quote or next-bar price rule fixed in advance, with explicit commissions, regulatory fees, spread costs, and a slippage function. The human's actual fills should be imported directly from the broker. To avoid giving the human an execution advantage that is really a data advantage, the fly simulator should use the same quote stream and should not fill at a midpoint unavailable to the human. If the human uses limit orders, missed fills and partial fills remain in the ledger rather than being retroactively converted to marketable trades.

Table 5

Proposed Prospective Matched Protocol

Element	Specification	Reason
Duration	252 regular U.S. market sessions; locked descriptive looks at 63 and 126	Covers varied regimes and prevents short-window storytelling.
Universe	SPY, QQQ, GLD, AAPL	Matches the prior fly study while remaining liquid.
Horizon	Intraday; flat by session close	Makes the human condition genuine day trading and removes overnight beta mismatch.
Market data	Common timestamped trades/quotes; prespecified 1- or 5-minute feature bars	Same information opportunity set.
Capital / leverage	Same starting equity, gross cap, per-symbol cap, and financing assumptions	Makes risk budgets comparable.
Execution	Human actual broker fills; fly conservative quote-based simulated fills with identical fee/slippage schedule	Prevents fantasy fills and keeps costs symmetric.
Strategy changes	None after start; emergencies recorded as protocol deviations	Separates a frozen test from research iteration.
Primary endpoint	Difference in mean net daily return, reported with paired day bootstrap / HAC-robust uncertainty	Same-day pairing controls shared market conditions.
Risk endpoints	Max drawdown, ES(5%), downside deviation, realized volatility, exposure, turnover	Separates return from risk budget.
Behavior endpoints	Post-loss and post-gain exposure change, size elasticity, trade-count response, rule deviations	Measures risk appetite operationally.
Multiplicity	One primary return endpoint; small prespecified secondary family; no winner by metric shopping	Protects against selection bias.

Note. The protocol is a research design, not a claim that the human or fly will be profitable. A preregistration and immutable configuration hash should be created before session 1.

Outcome Definitions

The primary daily account return should be based on marked end-of-day equity after all realized trading costs, with both participants flat at the close. With no external flows, the simple return is:

$$r_t = (V_t - V_{(t-1)}) / V_{(t-1)},$$

where V_t is end-of-day net liquidation value. If cash flows are permitted for operational reasons, a time-weighted return formulation should replace the simple ratio. Human and fly daily returns are paired by date. The primary estimand is the mean paired difference in net daily return, not whichever system happens to have the larger terminal wealth on one realized path.

Secondary performance measures should include annualized volatility, Sharpe ratio, Sortino ratio, cumulative log return, hit rate, payoff ratio, profit factor, maximum drawdown, 5% expected shortfall, skewness, excess kurtosis, and time under water. Sharpe inference must account for serial dependence rather than blindly applying square-root-of-time scaling (Lo, 2002). If many strategy variants are evaluated, the deflated Sharpe ratio or an explicit familywise procedure should be prespecified rather than used selectively after results are seen (Bailey & López de Prado, 2014; White, 2000).

Return structure is evaluated with daily and intraday autocorrelation, volatility clustering, P&L concentration by time of day, conditional results by realized-volatility regime, and gains/losses following prior gains/losses. The goal is not merely to ask which participant made more money, but to identify how each return process is generated.

Operational Measures of Human Risk Appetite

A single self-reported number should not be treated as risk appetite. The preregistered behavioral definition should be multivariate and observable. At each trade and at fixed intraday snapshots, record gross exposure, net exposure, largest position, notional trade size, leverage, turnover, number of open positions, distance to stop if a stop exists, and cumulative session P&L. The key state-dependence regression can be written as:

Risk_t = alpha + beta_loss I(P&L_(t-1) < 0)|P&L_(t-1)| + beta_gain I(P&L_(t-1) > 0)P&L_(t-1) + gamma X_t + epsilon_t,

where Risk_t can be gross exposure, trade size, or another prespecified risk proxy, and X_t contains market-volatility, time-of-day, and opportunity-set controls. A positive beta_loss is consistent with loss chasing or a break-even effect; a positive beta_gain is consistent with a house-money or self-attribution channel. These labels should remain interpretive, not causal, in a single-subject observational time series.

Additional human diagnostics should include the proportion of gains realized relative to losses, winner-versus-loser holding time, trade size after a sequence of wins or losses, frequency of discretionary overrides, and a brief pre-market/post-market risk-tolerance and affect survey. Any wearable physiology should be exploratory and separately consented; it is not required for the core economic comparison.

Table 6

Prespecified Outcome Hierarchy

Tier	Outcome	Interpretation
Primary	Mean paired human-minus-fly net daily return	Economic performance under matched days, costs, and risk limits.
Key secondary	Paired daily return volatility; max drawdown; ES(5%); downside deviation	Whether higher returns are purchased with greater realized risk.
Trading behavior	Exposure, turnover, trade count, holding time, concentration	How each participant uses the common risk budget.
Return structure	Autocorrelation, skewness, kurtosis, time-of-day P&L, volatility-regime P&L	Whether returns arise from stable or regime-dependent processes.
Human risk response	Post-loss and post-gain exposure/size coefficients	Operational evidence of path-dependent risk appetite.
Mechanism / fly	Ablations: no recurrence, random topology, calibrated strength-preserving null, native topology	Whether any machine edge is attributable to topology rather than generic memory or momentum.

Note. No single secondary metric should be allowed to overturn the primary endpoint after the fact. Positive secondary findings should be described as mechanism or risk evidence, not retroactively promoted to the main hypothesis.

Statistical Analysis Plan

Daily returns are naturally paired because both participants face the same calendar day. The main interval should therefore resample trading days in blocks or use a heteroskedasticity-and-autocorrelation-consistent estimator to preserve modest serial dependence. A moving-block bootstrap with a prespecified block length, alongside a Newey-West-style sensitivity analysis, is preferable to treating every intraday trade as an independent observation. Intraday trades are clustered within day; the day is the primary economic sampling unit.

For the single human trader, statistical uncertainty describes the observed trading process under the sampled market path; it does not create population inference over 'all humans.'

Similarly, computational seeds describe stochastic model construction, not biological specimens. A robust fly inference should cross market-day resampling with prespecified model seeds where possible, mirroring the crossed uncertainty logic used in the MaleCNS counting replication (Ames, 2026b).

The primary hypothesis should be directional only if direction is fixed before data collection. A neutral and defensible default is H0: the expected paired difference in net daily return is zero, with a two-sided interval. If the purpose is to test whether the fly architecture adds value over a human trader, the superiority margin and direction must be specified before session 1. Equivalence should not be claimed merely because a confidence interval includes zero; an equivalence margin and two one-sided tests must be prespecified.

Interpretation of Possible Outcomes

If the Human Wins

A human advantage would not prove biological superiority over connectome-derived computation. It could reflect discretionary information, execution skill, security selection, regime recognition, or simply a better match between human cognition and the intraday task. Mechanism analysis would ask whether the human advantage survives after matching exposure and turnover, whether it is concentrated in specific volatility regimes, and whether it is associated with stable skill or high-risk state dependence.

If the Fly Wins

A fly-derived advantage would be interesting but equally easy to overstate. The first comparison should be against conventional momentum, no-recurrence, random-topology, and calibrated graph nulls—not against the human alone. If all recurrent models beat the human equally, the result would support an automated momentum policy, not fly anatomy. If native topology beats matched nulls prospectively, the claim can narrow to incremental value from that engineered topology under the tested dynamics. It still would not show that a living fly can trade.

If Neither Demonstrates Positive Net Economics

This may be the most scientifically useful result. The prior fly paper already showed that an elaborate recurrent architecture can be technically successful and economically unconvincing. Human day-trading studies show the same separation from the other direction: sophisticated, persistent, and psychologically engaged trading often fails after costs. A null or mutually negative prospective result would therefore reinforce a general lesson: complexity, activity, and adaptive behavior are not themselves sources of edge.

Discussion

The two supplied papers form a coherent methodological sequence. The temporal-counting work established that connectome-derived topology can alter artificial recurrent computation under fixed assumptions, but also showed that simple task structure and engineering choices can dominate the narrative. The momentum-trading work then tested whether that unusual topology translated into economic value and found no demonstrated advantage. The proposed human comparison is valuable precisely because it should not restart the narrative at a lower evidentiary standard.

The human literature provides a powerful baseline prior. Most day traders do not earn durable net profits. Yet that statement should not be converted into 'humans are bad traders' in a way that erases heterogeneity. Barber et al. (2014) demonstrate a narrow but real group with persistent abnormal profits. That upper tail is the relevant scientific reason to test a preidentified human prospectively rather than selecting a winner retrospectively. If the trader is chosen because their historical record already looks extraordinary, the prospective phase becomes essential; otherwise the study simply compares a selected human backtest with a selected machine backtest.

Risk appetite is where the comparison is most conceptually revealing. A fly-derived reservoir may possess memory, but it has no valence. It does not care whether yesterday's loss was humiliating or whether a morning gain feels like house money. Human traders can change their risk budget because of outcome feedback even when the market's opportunity set has not changed. That dynamic can create both adaptability and error. From an engineering perspective, the human is a controller whose internal state is only partially observed; the reservoir is a controller whose state equation is explicit. This makes the machine easier to audit and the human potentially richer but harder to identify.

The return series consequently carries different information. For the reservoir, autocorrelation or tail events primarily reflect market structure, recurrent memory, gating, and exposure rules. For the human, the same return statistics can also encode psychological feedback, attention, fatigue, confidence, and rule changes. A matched study should therefore treat orders and exposure decisions as primary data, not merely intermediate steps on the way to cumulative P&L.

Limitations

The present paper has several limitations. First, it contains no new human trading ledger and therefore cannot answer who is more profitable. Second, the existing fly financial evidence is daily, while day trading is intraday. Third, the human literature spans different countries, periods, products, cost structures, and definitions; the numerical success rates shown together are not directly comparable. Fourth, the fly architecture is only loosely biological: topology is connectome-derived, but state dynamics, signs, leak, scaling, features, and readouts are engineering choices. Fifth, a single prospective human trader would not represent the population of humans; it would be a case study with dense repeated measures. Sixth, a one-year market sample can still be regime specific. The strongest program would eventually replicate across additional humans, additional periods, and independently frozen fly configurations.

There is also a subtle selection problem. If the human subject is chosen because they are already known to be unusually successful, the study estimates the performance of a selected trader, not the expected performance of a new day trader. That is acceptable if stated explicitly. What is not acceptable is to generalize that selected case to human traders broadly without a defined sampling frame.

Conclusion

The existing evidence does not support a claim that fly-derived reservoir trading outperforms human day trading, nor does it support the reverse. The reason is not statistical ambiguity alone; it is a design mismatch. The fly financial study is a retrospective daily momentum allocation experiment. Human day trading is intraday, execution intensive, behaviorally adaptive, and strongly affected by transaction costs. Direct ranking would be apples versus caffeinated oranges.

The defensible comparison is structural. The fly-derived system currently has a frozen, engineered risk policy whose returns depend on price/volatility inputs, recurrent memory, gating, exposure, and costs. Human day traders exhibit a broad and usually unfavorable return distribution with a small persistent-skill tail, and their risk taking can change after gains, losses, emotional reactions, confidence updates, and salient attention. In that sense, the human has risk appetite; the reservoir has risk parameters.

The highest-value next experiment is therefore prospective and synchronized: rerun the fly system intraday, put one preidentified human trader under the same assets, capital, leverage, session, and cost rules for 252 sessions, freeze both protocols before the first trade, and compare paired net returns while measuring how each uses risk. If the result survives conventional momentum controls, random and strength-preserving graph nulls, transaction costs, and locked aggregation rules, it will answer a substantially stronger question than another retrospective backtest. Until then, the correct conclusion is disciplined uncertainty, not a story about whether a fly can beat a human trader.

Data and Research Transparency

All fly-system numerical values in this manuscript were transcribed from the two supplied manuscripts and cross-checked against their reported tables and figures. Human-study numerical values were taken from the cited primary publications or their official abstracts/repository versions. No trader-level human dataset was downloaded or reanalyzed for this paper, and no synthetic human return series was generated. The figures in this manuscript are visualizations of reported summary statistics or conceptual diagrams. A future empirical paper should archive the matched market stream, human order/fill ledger, fly decisions, configuration hashes, and analysis code before unblinding final outcomes.

References

Ames, O. (2026a). *Fly-derived reservoir computing for momentum trading: Retrospective evidence, engineering strengths, and empirical failures* [Unpublished manuscript]. Ames Investment Systems.

Ames, O. (2026b). *Temporal counting in MaleCNS-derived artificial recurrent networks: An independent replication, memory-capacity extension, and synchronized blackjack simulation* [Unpublished manuscript]. Arizona State University.

Bailey, D. H., & López de Prado, M. (2014). The deflated Sharpe ratio: Correcting for selection bias, backtest overfitting, and non-normality. *The Journal of Portfolio Management*, *40*(5), 94-107. https://doi.org/10.3905/jpm.2014.40.5.094

Barber, B. M., Lee, Y.-T., Liu, Y.-J., & Odean, T. (2014). The cross-section of speculator skill: Evidence from day trading. *Journal of Financial Markets, 18*, 1-24. https://doi.org/10.1016/j.finmar.2013.05.006

Barber, B. M., Lee, Y.-T., Liu, Y.-J., Odean, T., & Zhang, K. (2020). Learning, fast or slow. *The Review of Asset Pricing Studies*, *10*(1), 61-93. https://doi.org/10.1093/rapstu/raz006

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance, 55*(2), 773-806. https://doi.org/10.1111/0022-1082.00226

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics, 116*(1), 261-292. https://doi.org/10.1162/003355301556400

Ben-David, I., Birru, J., & Prokopenya, V. (2018). Uninformative feedback and risk taking: Evidence from retail Forex trading. *Review of Finance*, *22*(6), 2009-2036. https://doi.org/10.1093/rof/rfy022

Berg, S., Beckett, I. R., Costa, M., Schlegel, P., Januszewski, M., Marin, E. C., Nern, A., Preibisch, S., Qiu, W., Takemura, S., Fragniere, A. M. C., Champion, A. S., Adjavon, D.-Y., Cook, M., Gkantia, M., Hayworth, K. J., Huang, G. B., Kampf, F., Katz, W. T., et al. (2026). Sexual dimorphism in the complete Drosophila male central nervous system connectome. *Cell*, *189*(18), 5504-5526.e15. https://doi.org/10.1016/j.cell.2026.08.015

Chague, F., De-Losso, R., & Giovannetti, B. (2020). *Day trading for a living?* (Working Paper 525). Fundação Getulio Vargas, Escola de Economia de São Paulo. https://hdl.handle.net/10438/28801

Chague, F., & Giovannetti, B. (2025). The COVID-19 and day-trade pandemics in Brazil. *Brazilian Review of Finance*, *23*, e202515. https://doi.org/10.12660/rbfin.v23n1.2025.94291

Chague, F., Giovannetti, B., & Paiva, G. (2026). Familiarity breeds day trade. *Journal of Banking & Finance*, *185*, 107651. https://doi.org/10.1016/j.jbankfin.2026.107651

Coval, J. D., & Shumway, T. (2005). Do behavioral biases affect prices? *The Journal of Finance*, *60*(1), 1-34. https://doi.org/10.1111/j.1540-6261.2005.00723.x

Dorn, D., & Huberman, G. (2005). Talk and action: What individual investors say and what they do. *Review of Finance*, *9*(4), 437-481. https://doi.org/10.1007/s10679-005-4997-z

Grinblatt, M., & Keloharju, M. (2009). Sensation seeking, overconfidence, and trading activity. *The Journal of Finance*, *64*(2), 549-578. https://doi.org/10.1111/j.1540-6261.2009.01443.x

Huang, Y.-C., & Chan, S.-H. (2014). The house money and break-even effects for different types of traders: Evidence from Taiwan futures markets. *Pacific-Basin Finance Journal*, *26*, 1-13. https://doi.org/10.1016/j.pacfin.2013.10.008

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65-91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Jordan, D. J., & Diltz, J. D. (2003). The profitability of day traders. *Financial Analysts Journal*, *59*(6), 85-94. https://doi.org/10.2469/faj.v59.n6.2578

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263-291. https://doi.org/10.2307/1914185

Kuhnen, C. M., & Knutson, B. (2005). The neural basis of financial risk taking. *Neuron*, *47*(5), 763-770. https://doi.org/10.1016/j.neuron.2005.08.008

Kumar, A. (2009). Who gambles in the stock market? *The Journal of Finance*, *64*(4), 1889-1933. https://doi.org/10.1111/j.1540-6261.2009.01483.x

Lo, A. W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, *58*(4), 36-52. https://doi.org/10.2469/faj.v58.n4.2453

Lo, A. W., & Repin, D. V. (2002). The psychophysiology of real-time financial risk processing. *Journal of Cognitive Neuroscience*, *14*(3), 323-339. https://doi.org/10.1162/089892902317361877

Lo, A. W., Repin, D. V., & Steenbarger, B. N. (2005). Fear and greed in financial markets: A clinical study of day-traders. *American Economic Review*, *95*(2), 352-359. https://doi.org/10.1257/000282805774670095

Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics*, *104*(2), 228-250. https://doi.org/10.1016/j.jfineco.2011.11.003

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703-708. https://doi.org/10.2307/1913610

Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of Finance*, *53*(5), 1775-1798. https://doi.org/10.1111/0022-1082.00072

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297-323. https://doi.org/10.1007/BF00122574

White, H. (2000). A reality check for data snooping. *Econometrica*, *68*(5), 1097-1126. https://doi.org/10.1111/1468-0262.00152

Appendix A

Core Metric Definitions for the Prospective Study

The following definitions are recommended to prevent metric drift after outcomes are observed.

They are not new estimates from the present paper.

Table A1

Metric Definitions

Metric	Definition	Use / caveat
Net daily return	(end-of-day net liquidation value - prior end-of-day value) / prior end-of-day value	Primary economic unit; after fees and slippage.
Annualized volatility	sqrt(252) x SD(daily returns)	Descriptive scale; report nonannualized SD too.
Sharpe ratio	mean(excess daily return) / SD(excess daily return), with serial-dependence-aware inference	Do not rely on naive iid standard errors.
Sortino ratio	mean(return - target) / downside deviation	Target fixed in advance; unstable with few downside observations.
Maximum drawdown	max peak-to-subsequent-trough loss in cumulative equity	Path dependent; not a timing score by itself.
Expected shortfall (5%)	mean return among observations at or below the 5th percentile	Tail-loss severity.
Turnover	sum absolute notional changes / average equity	Direct cost and activity proxy.
Gross exposure	sum absolute market value of positions / equity	Primary operational risk-budget measure.
Net exposure	signed market value / equity	Directional market beta proxy.
Post-loss risk coefficient	slope linking prior negative P&L magnitude to subsequent exposure/size	Operational break-even/loss-chasing measure.
Post-gain risk coefficient	slope linking prior positive P&L to subsequent exposure/size	Operational house-money/self-attribution measure.

Note. All definitions should be frozen in the preregistration. Intraday observations are clustered within day for inferential purposes.

Appendix B

Minimum Preregistration Checklist

- Human participant identity or sampling rule fixed before session 1; no replacement for poor early performance.
- Fly model checkpoints, graph families, seeds, feature code, and hyperparameters hashed and frozen.
- Tradable universe, market-data vendor, decision interval, session hours, and corporate-action handling fixed.
- Starting equity, leverage cap, per-symbol limits, shorting rules, and forced end-of-day liquidation fixed.
- Commission, fee, spread, slippage, and simulated-fill rules fixed and testable on historical fixtures.
- Primary return endpoint and exact aggregation rule fixed; secondary family enumerated.
- Block/bootstrap or HAC inference method, block length rule, and multiplicity procedure fixed.
- Human discretionary rule changes prohibited or preclassified as protocol deviations; emergency criteria defined.
- No interim model tuning after 63- or 126-session descriptive checkpoints.
- Raw orders, cancels, fills, positions, P&L, market states, and model decisions archived with synchronized timestamps.
- Final analysis executed from an immutable script against an untouched final dataset snapshot.