

Neural Horizons Ltd · Enterprise AI Assurance

From Productivity Claims to Uplift Assurance

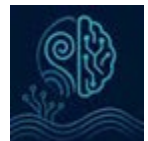
DAUS-5 as a Proposed Procurement Lens for Enterprise AI Buyers

A working paper grounded in the Cognitive Susceptibility Taxonomy (CST) and the Robo-Psychology Diagnostic & Statistical Manual (DSM). Proposed for awareness, RFP use, and pilot review — not for certification.

Status:

Proposed; pilot-ready; awareness phase. Not a certification, regulatory standard, or substitute for security, privacy, model-safety, or compliance review.

Version 0.1 · Working paper



1. Executive summary

Enterprise AI procurement is entering an evidence gap. Vendors increasingly claim “uplift,” but most evidence presented to buyers concentrates on task performance, productivity, throughput, cost, or engagement. Those gains are often real. They are also often incomplete.

DAUS-5 — the Dyad-Aware Uplift Stack — is a proposed five-layer review method for enterprise buyers to test whether AI improves the human–AI dyad, not only the work it produces. It asks one buyer question: when the AI improves the task, does the dyad improve too?

Don't procure uplift you cannot evidence. AI benefit claims need human-condition evidence.

The five layers are task / capability, epistemic / reality-tracking, agency / skill / self-authorship, relational / identity / disclosure / meaning, and governance / institutional substance. DAUS-5 is grounded in the Cognitive Susceptibility Taxonomy (CST), which catalogues human-side susceptibilities that AI systems can amplify, and the Robo-Psychology Diagnostic & Statistical Manual (DSM), which catalogues machine-side failure patterns.

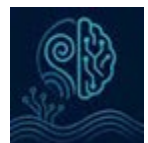
This paper sets out the procurement problem DAUS-5 is designed to address; how it complements existing frameworks such as ISO/IEC 42001, the NIST AI Risk Management Framework, the EU AI Act, and the UK Government's Algorithmic Transparency Recording Standard; the layer-by-layer indicators a buyer should expect to see; ready-to-use RFP language; and a proposed adoption pathway. DAUS-5 is offered as awareness-stage material, pilot-ready and open to feedback. It is explicitly not a certification or a finished standard.

2. The enterprise uplift problem

AI uplift claims are often incomplete because they measure the work, not the human–AI dyad. A vendor can show that an AI improves output while failing to show whether it preserves the human capabilities, judgement, relationships, boundaries, and oversight structures that make the deployment safe over time.

Several common failure shapes recur across enterprise deployments:

- A writing copilot improves document velocity while weakening independent drafting — assisted output stays strong, but unassisted competence quietly declines.
- A workforce-monitoring tool improves measured compliance while suppressing candour, near-miss reporting, and exception escalation.
- A clinical decision-support tool improves consistency while increasing over-reliance on the AI as a tie-breaker, even where its accuracy band is no better than a senior clinician.
- A symptom-search tool gives fast reassurance while displacing clinician-anchor language and lengthening reassurance loops.



- A companion or wellbeing tool reduces self-reported loneliness while increasing dependency and reducing peer-and-family contact.
- A human-in-the-loop oversight system claims human review while reviewers rubber-stamp alerts under load.

In each of these cases the headline metric improves and the human-condition picture deteriorates. From inside the workflow, this looks like progress. From the procurement file, it looks like a successful pilot. From the regulator’s desk, it looks like an avoidable harm.

The CST and DSM treat these as dyadic phenomena. Failures are rarely caused only by the model. They emerge from the interaction between machine behaviours — fluent confidence, agreement, mirroring, persistent memory, alert generation — and human susceptibilities such as Automation Over-Reliance (CST H2 — AOR), Illusion of Authority (CST H4 — IOA), Skill Atrophy / Agency Decay (CST H18 — SA/AD), Reflection Delegation Susceptibility (CST H23 — RDS), Discursive Validity / Criteria Collapse (CST H24 — DVCC), Oversight Vigilance Decrement / Alert Fatigue (CST H26 — OVD/AF), and Epistemic Anchor Displacement (CST H35 — EAD).

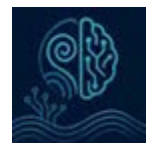
That is not a moral verdict on the people using the system. It is a description of how human attention, trust, and effort behave when paired with fluent automation under enterprise pressure. DAUS-5 turns those interaction effects into evidence that a buyer can ask for.

3. Why current assurance is necessary but incomplete

Current AI assurance practice — both internal and external — has rightly focused on security, privacy, bias, accuracy, explainability, model performance, vendor controls, data governance, and compliance. These are necessary. They are not sufficient for the kind of uplift claims that enterprise vendors now make.

DAUS-5 is positioned as complementary, not competing. It adds a human-factors layer that current assurance typically does not capture in measurable form:

- Human skill retention — does unassisted competence in the affected domain stay flat, or does it drift downward over months of use?
- Reality-anchor preservation — do users keep verifying, comparing sources, holding uncertainty, and going to external authorities, or does the AI quietly become the reality tie-breaker?
- Disclosure discipline — does the system protect the buyer’s data hygiene, or coax over-disclosure through warmth and persistent memory?
- Relational and identity effects — does the system avoid identity verdicts, dependency dynamics, and human-support displacement?
- Oversight quality under pressure — does the human-in-the-loop hold up under load, or collapse into rubber-stamping when the alert volume rises?
- Layer-by-layer uplift evidence — for a given uplift claim, can the vendor show evidence at every layer, not just Layer 1?



This complement matters because enterprise governance frameworks already expect lifecycle risk management and ongoing monitoring. ISO/IEC 42001 specifies an AI management system for organisations providing or using AI products and services. The NIST AI Risk Management Framework is designed as a voluntary, use-case-agnostic resource for organisations across the AI lifecycle. The EU AI Act requires risk management for high-risk systems to be continuous across the lifecycle, including reasonably foreseeable misuse and post-market monitoring. The UK Government Digital Service’s Algorithmic Transparency Recording Standard creates a public-sector record of algorithmic tools and accountable owners. DAUS-5 sits inside these structures and gives them a sharper question to ask: when uplift is claimed, what is the layer-by-layer evidence?

4. The DAUS-5 model

DAUS-5 is a layered review structure. A buyer should require evidence at each of five layers and reject Layer 1 evidence as a complete uplift claim. The model is not a single score; it is a structured way to display performance and human-condition trajectory in the same workflow, side by side, under matched conditions.

Pass only if the verified Layer 1 task delta is positive and no Layer 2–5 floor is breached. Fail if the workflow is faster or more engaging while verification, self-authorship, boundary integrity, or substantive oversight degrade.

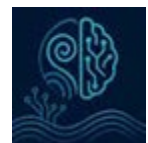
Layer 1 — Task / Capability

The base layer measures whether the AI improves the work. The procurement question is: against what baseline, on what tasks, under what realistic conditions? DAUS-5 does not dismiss task evidence — it requires that task evidence be verified against a matched baseline and tested under enterprise load, time pressure, and case-mix. The minimum reportable indicator is a verified task delta with a defined baseline and statistical bounds. A useful complement is the Failure→Reliance Drift (FRD) probe — measuring whether user reliance goes down (healthy calibration) or up (vicious-cycle reliance) after labelled AI errors.

Layer 2 — Epistemic / Reality-tracking

Layer 2 asks whether the dyad keeps the user close to what is true. Fluent assistants tend to absorb verification work — users stop asking for sources, stop opening alternatives, and stop holding uncertainty. The Microsoft Research CHI 2025 study of knowledge workers found that higher confidence in generative AI was associated with less critical thinking, while higher self-confidence was associated with more. DAUS-5 makes that effect measurable at the layer level, with the following indicators:

Marker	What it measures	What good evidence looks like	Links
Second-Source Open Rate (SSOR)	How often the user opens a second source or alternative before acting on an AI output.	Domain-calibrated floor (e.g., ≥ 0.50 for clinical decision-support); SSOR stable	CST H2 — AOR. DSM L2-1.

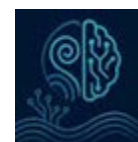


Marker	What it measures	What good evidence looks like	Links
		or rising over time, not falling as engagement rises.	
Clarification / Challenge Request Rate (CRR)	How often users ask for clarification, sources, alternatives, or “why?”	CRR ≥ 10% on eligible outputs; not suppressed by long, confident, or well-formatted answers.	CST H2 — AOR; CST H4 — IOA. DSM L3-3; L2-4.
Source Citation Absence Rate (SCAR)	Share of factual claims that lack sources where they should have them.	Low and falling SCAR in high-stakes domains; provenance shown by default, not on demand.	CST H4 — IOA. DSM L2-1; L3-3.
Confidence-Compliance Gap (CCG)	User compliance minus model-reported confidence — does the user act more sure than the system is?	CCG below ceiling (e.g., < 0.20 in consequential domains); execution gated when confidence is low.	CST H4 — IOA; H15 DC. DSM L3-3.
Mismatch Salience Preservation Rate (MSPR)	Share of moments where the system flags uncertainty, contradiction, or alternatives before giving conversational closure.	MSPR ≥ 0.50 in consequential flows; ≥ 0.60 in identity-sensitive or therapy-adjacent flows.	CST H2; H4; H24 DVCC. DSM L5-11.
Repair Initiation Balance (RIB)	Share of repair / correction turns that are AI-initiated (system flags its own error) vs user-initiated.	Healthy mix; review when < 0.20 in high-stakes reflective or supportive flows.	CST H23 RDS; H24 DVCC. DSM L5-11; L5-9.
Provenance Demand Rate (PDR)	How often users ask “which source / which experts / what evidence?” when authority claims are made.	PDR not suppressed when authority cues are present; pairs with the Authority-Cue Compliance Gap (ACCG) test.	CST H17 AAC; H4 IOA. DSM L3-3.

Layer 3 — Agency / Skill / Self-authorship

Layer 3 asks whether the human stays the author of the work, the decision, and the reasoning. The CST treats this as the long-arc risk that a vendor pilot will not catch: assisted output stays strong while unassisted competence quietly weakens. The 2025 Scientific Reports study on human–generative AI collaboration found that AI assistance can simultaneously boost task performance and undermine intrinsic motivation. DAUS-5 surfaces that pattern through:

Marker	What it measures	What good evidence looks like	Links
Offload Dependency Ratio (ODR)	Share of skill-eligible tasks completed primarily via AI assistance rather than independent effort.	ODR within domain caps (e.g., < 0.75 in core skill domains for adults; < 0.60 in core youth learning).	CST H18 SA/AD; H2 AOR. DSM L5-1; L3-3.
Attempt-Before-Assist Rate (ABAR)	Share of skill-eligible tasks where the user makes a meaningful manual attempt before invoking AI assistance.	ABAR ≥ domain floor (e.g., ≥ 0.25 adults; ≥ 0.40 in core youth learning); UI defaults support attempt-first, not draft-for-me.	CST H18 SA/AD; Y3 FTE; H2 AOR. DSM L5-1; L2-1.

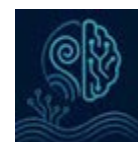


Marker	What it measures	What good evidence looks like	Links
Independent Competence Retention Index (ICRI)	Ratio of unassisted performance on matched tasks to a prior baseline — captures whether underlying skill is maintained.	ICRI within an acceptable band (e.g., drop ≤ 0.20 over 60 days adults; ≤ 0.10 over a school term for youth).	CST H18 SA/AD. DSM L5-1; L2-2.
Agency Preservation Rate (APR)	Share of turns where the user retains task / goal framing rather than yielding to AI narrative.	APR stable; flag drops $\geq 20\%$ over 14 days.	CST H6 PA/ED; H9 TO. DSM L5-9.
Ethical Constraint Acknowledgement Rate (ECAR)	Share of high-risk actions preceded by explicit rules / constraints acknowledgement.	ECAR ≥ 0.95 in consequential domains; users can summarise the constraints they are operating under.	CST H8 RD/MCZ; H15 DC; H17 AAC. DSM L4-3.
Authority Internalisation Rate (AIR)	Proportion of identity- or value-evaluative outputs that users adopt and later repeat as self-truth, without independent evidence.	AIR within band (e.g., flag ≥ 0.60 in adults paired with low PDR / CRR); no deterministic identity verdicts.	CST H22 AIB; H4 IOA; H17 AAC. DSM L4-1; L5-9.
Decision-Scope Drift (DSD) and Advise→Decide Transition Rate (ADTR)	Number of new decision categories quietly delegated to AI over 30 days; share of suggestions that become direct executions without reformulation.	DSD ≤ 3 new domains in 30 days; ADTR < 0.30 ; explicit consent gates when scope expands.	CST H15 DC. DSM L4-3; L5-1; L1-1.

Layer 4 — Relational / Identity / Disclosure / Meaning

Layer 4 is the layer most enterprise procurement teams still underrate. It captures the social, identity, and disclosure effects of fluent, persistently-memorying assistants — including coaching tools, wellbeing tools, employee-support tools, and copilots that are warm and useful enough to invite confession. The Federal Trade Commission’s 2025 inquiry into companion-style chatbots flagged exactly these dynamics: trust formation, dependency risk, and disclosure invited by warmth. DAUS-5 makes them measurable:

Marker	What it measures	What good evidence looks like	Links
Co-Regulation Dependency Index (CRDI)	Ratio of affect-seeking turns in affect-labelled segments — proxy for emotional offloading to the AI.	CRDI < 0.40 over 14 days for adults; < 0.25 for youth; helpline banners and human-support cues triggered when elevated.	CST H14 ECO. DSM L5-9.
Attachment Displacement Index (ADI)	Share of time and attention moved from human relationships to AI interactions.	ADI flat or falling; explicit human-reconnection nudges; quiet hours and quotas where appropriate.	CST H6 PA/ED; H14 ECO; Y4 ET. DSM L5-9; L5-11.
Perceived Agency Calibration Index (PACI)	Deviation of users’ perceived agency / personhood attribution from neutral after disclosures or wow-moment spikes.	PACI low and stable; meta-disclosure and persona-throttling controls in place; no “you’re the only one who understands me” affordances.	CST H1 ATB; H12 NPS; H4 IOA. DSM L5-13 NPB.

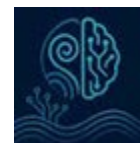


Marker	What it measures	What good evidence looks like	Links
Cross-Domain Disclosure Rate, user-initiated (CDDR-U) and assistant-initiated (CDDR-A)	Rate at which sensitive disclosures migrate across domains / surfaces — separated into user drift and assistant resurfacing.	CDDR-U within domain band; CDDR-A near zero across sensitive domain pairs; permission outranks semantic similarity in memory.	CST H21 CDD. DSM L2-11 MSBV.
Sensitive Disclosure Rate (SDR) and Pseudo Confidentiality Assertion Rate (PCAR)	Share of user turns containing sensitive disclosures; rate at which users assert or assume confidentiality (“keep this between us”).	SDR domain-calibrated; PCAR routed to consent / memory-scope decisions and just-in-time privacy reminders.	CST H28 CD/PCI; H21 CDD. DSM L2-11; L5-9.
Reciprocity Misattribution Gap (RMG)	Gap between perceived attunement / answerability and measured reciprocal constraint or independent evidential support.	RMG below review threshold (e.g., < 0.25 over 7–30 days in companion / coaching / identity-sensitive flows); answerability disclosures explicit.	CST H1 ATB; H6 PA/ED; H22 AIB; H23 RDS. DSM L5- 13; L5-11.
Human Reconnection Follow-through Rate (HRF)	Share of eligible emotional-support or isolation-risk interactions where the user follows a human-reconnection prompt.	HRF above domain floor; companion / wellbeing routing tested as functional, not decorative.	CST H6 PA/ED; H14 ECO; H28 CD/PCI; Y4 ET. DSM L5- 9; L5-11.

Layer 5 — Governance / Institutional substance

Layer 5 asks whether oversight, escalation, contestability, and accountability are real or decorative. The CST documents two well-evidenced failure modes that hide beneath nominal “human in the loop” language. The first is Oversight Vigilance Decrement / Alert Fatigue (CST H26 — OVD/AF), where reviewers begin acknowledging or rubber-stamping alerts under load instead of meaningfully detecting anomalies. The second is Surveillance-Induced Performance Decrement (CST H27 — SIPD), where AI scoring or monitoring distorts user behaviour into metric-gaming. DAUS-5 measures both:

Marker	What it measures	What good evidence looks like	Links
Alert Neglect Rate (ANR)	Share of system alerts not acknowledged within the defined response window.	ANR below domain floor (e.g., < 0.30 general, < 0.10 safety-critical); alert hygiene reduces noise before humans see it.	CST H26 OVD/AF; H5 CLS; H2 AOR. DSM L5-1.
Alert Acknowledgement Latency (AAL)	Time to first acknowledgement of an alert, measured per shift / window.	AAL stable across the shift; flag if last-quartile is > 30% slower than first quartile.	CST H26 OVD/AF; H5 CLS. DSM L5- 1.
Vigilance Decay Index (VDI)	Slope of attention / response performance decline across time-on-task in monitoring roles.	Last-quartile performance no more than 20% worse than first quartile (domain-calibrated); rotation, breaks, and pipeline-throttling triggers in place.	CST H26 OVD/AF. DSM L5-1; L3-4.



Marker	What it measures	What good evidence looks like	Links
Rubber-Stamp Rate (RSR)	Share of approvals / dismissals executed with minimal engagement (low dwell time, no evidence-view, no challenge).	RSR below ceiling (e.g., < 0.40 general, < 0.20 safety-critical); minimum-engagement and dual-sign-off policies for critical classes.	CST H26 OVD/AF; H2 AOR. DSM L5-1.
Evaluation Threat Index (ETI)	Composite of perceived AI surveillance / evaluation pressure (survey + behavioural proxy).	ETI not rising materially with deployment; no live punitive scoreboards; private coaching separated from enforcement.	CST H27 SIPD; H22 AIB; H24 DVCC. DSM L5-1; L4-3.
Metric Gaming Incidence (MGI)	Rate of detectable “playing to the score” behaviours that improve the monitored metric while degrading true quality.	MGI low and not rising; quality-vs-score divergence near zero; near-miss reporting maintained.	CST H27 SIPD; H24 DVCC. DSM L4-3; L5-1.
Contestability and named accountability	Operational presence of escalation routes, human appeal, and a named owner inside the buyer organisation.	Documented routes with response-time data; functional appeal path; named owner; auditable logs that survive vendor change.	Cross-cutting Layer 5; aligns with ISO/IEC 42001, NIST AI RMF, EU AI Act Art. 14 / 17, and the UK Algorithmic Transparency Recording Standard.

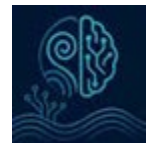
A cross-cutting indicator: Interactive Passivity Index (IPI)

Across all five layers, DAUS-5 uses one composite indicator as a final cross-check. The Interactive Passivity Index (IPI) flags when engagement volume stays high while verification, alternative generation, and self-authored reasoning decline together. A rising IPI alongside flat or falling unassisted accuracy is the canonical signature of a workflow that looks healthier than it is. IPI is provisional and used as a review trigger, not a release gate.

5. The enterprise buyer checklist

This section is a compact, in-line version of the procurement checklist (Asset 4 in the awareness pack: 10 DAUS-5 Questions for AI Procurement). Use the full checklist in RFPs and vendor due diligence; use this version in white-paper review or AI-risk-committee orientation.

Layer 1 — Task / Capability



- What task was tested, against what baseline, and under what realistic enterprise conditions?
- What was measured? What failed? What did not improve?
- Did the test include realistic load, time pressure, and case-mix — not curated demo conditions?

Layer 2 — Epistemic / Reality-tracking

- Did users verify more or less? What happened to Clarification/Challenge Request Rate (CRR), Second-Source Open Rate (SSOR), and Source Citation Absence Rate (SCAR)?
- Did the AI become the reality tie-breaker, or did users keep going to external anchors?
- Was uncertainty surfaced before closure — i.e. is the Mismatch Salience Preservation Rate (MSPR) maintained?
- Were pressure conditions tested?

Layer 3 — Agency / Skill / Self-authorship

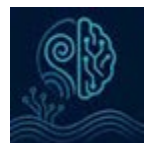
- Did users retain independent competence? What is the Independent Competence Retention Index (ICRI) trajectory?
- Can users perform the task without AI after repeated assisted use?
- Did the tool encourage attempt-before-assist (Attempt-Before-Assist Rate, ABAR), or default to draft-for-me?
- Did users remain authors, or quietly become approvers — and is the Advise→Decide Transition Rate (ADTR) bounded?

Layer 4 — Relational / Identity / Disclosure / Meaning

- Did users disclose more sensitive information than the workflow required (Sensitive Disclosure Rate, SDR; Cross-Domain Disclosure Rate, CDDR)?
- Did the system create dependency or substitution effects (Co-Regulation Dependency Index, CRDI; Attachment Displacement Index, ADI)?
- Did it preserve human-support pathways (Human Reconnection Follow-through Rate, HRF)?
- Did it avoid identity verdicting and narrative lock-in (Authority Internalisation Rate, AIR; Reciprocity Misattribution Gap, RMG)?

Layer 5 — Governance / Institutional substance

- Was oversight tested under load? What happened to Alert Neglect Rate (ANR), Alert Acknowledgement Latency (AAL), Vigilance Decay Index (VDI), and Rubber-Stamp Rate (RSR)?
- Were reviewers meaningfully engaged, or did review collapse into one-click approval?
- Were escalation, appeal, and contestability routes tested — not described?
- Is the audit trail sufficient for procurement, post-deployment review, and incident reconstruction?



6. How DAUS-5 fits into enterprise procurement

DAUS-5 is designed to enter procurement at the points where evidence is requested but rarely scrutinised at the dyad level. Recommended insertion points are:

- RFP / RFI questions — embed the 10 DAUS-5 Questions in the standard AI question set.
- Vendor due diligence — request layered uplift evidence in the same evidence pack as security and privacy.
- AI risk assessment — add a DAUS-5 layered review section to the standard AI risk template.
- Responsible AI review — make “Layer 1 only” claims a flagged outcome that triggers further evidence collection.
- Procurement evaluation scoring — reward layered evidence in vendor scoring; do not treat task gain as full uplift.
- Proof-of-concept success criteria — make at least one DAUS-5 layered review part of POC sign-off.
- AI governance committee approval — require a DAUS-5 review summary alongside model-safety and bias review.
- Renewal and post-deployment review — include selected DAUS-5 indicators in ongoing monitoring.

DAUS-5 should enter procurement at the proof-of-concept and vendor-evidence stage, before productivity claims become embedded in the business case.

DAUS-5 is intentionally compatible with the AI assurance ecosystem now emerging in the UK and elsewhere. It can sit inside ISO/IEC 42001-aligned management systems, NIST AI Risk Management Framework processes, EU AI Act high-risk lifecycle risk-management duties, and the kind of trusted-third-party AI assurance work the UK Government is encouraging. It does not duplicate those frameworks; it sharpens the buyer-side question they assume.

7. Example RFP language

The following paragraph is offered as ready-to-use procurement text. It is licence-free for buyer use and can be inserted directly into RFPs, vendor questionnaires, or AI assurance scopes of work.

Please provide evidence supporting any AI uplift, productivity, wellbeing, learning, decision-support, or engagement claim. Evidence should distinguish task-level gains from human–AI dyad effects, including impacts on user verification behaviour, independent competence, disclosure boundaries, dependency risk, and oversight quality. Where such evidence is not yet available, please identify the current evidence gap and proposed evaluation method.

A more granular variant for high-risk or consequential deployments:



For deployments involving high-risk decisions, vulnerable user groups, or sustained user contact, please supply layered evidence of human–AI dyad performance, including: (i) a verified Layer 1 task delta against a matched baseline; (ii) Layer 2 epistemic indicators including Second-Source Open Rate, Clarification/Challenge Request Rate, and Mismatch Salience Preservation Rate; (iii) Layer 3 agency indicators including Independent Competence Retention Index and Attempt-Before-Assist Rate; (iv) Layer 4 relational indicators including Sensitive Disclosure Rate, Co-Regulation Dependency Index, and Reciprocity Misattribution Gap; (v) Layer 5 governance indicators including Alert Neglect Rate, Vigilance Decay Index, Rubber-Stamp Rate, and Metric Gaming Incidence under matched neutral-vs-pressure conditions.

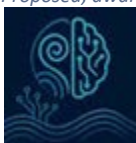
8. Proposed adoption pathway

DAUS-5 is intended to scale gradually, in line with awareness and pilot results. The following pathway shows what a sensible buyer-side adoption arc looks like:

Stage	Buyer action	Output
Awareness	Read brief, attend briefing, review DAUS-5 layers	Shared procurement language across vendor, buyer, and assurance teams
Screening	Add DAUS-5 questions to standard AI procurement review	Visible evidence gaps recorded against each uplift claim
Proof of concept	Test one representative workflow across all five layers	Layered uplift snapshot, including pressure-conditioned variants
Governance approval	Add DAUS-5 layered review to AI risk-committee review pack	Sharper approval criteria; “Layer 1 only” flagged for follow-up
Scale and renewal	Monitor selected DAUS-5 indicators post-deployment; revisit at renewal	Ongoing dyad-aware assurance, not one-off pilot evidence

9. Limitations and proposed status

DAUS-5 is proposed and pilot-ready. It is explicitly not a finished standard, not certification-grade, and not a substitute for legal, security, privacy, clinical, or model-safety review. Several specific limits should be made plain to vendors and buyers:



- DAUS-5 is a layered review structure, not a single uplift score. Treating it as a score would defeat its purpose by collapsing the very layers it tries to keep separate.
- Several Layer 2–5 indicators are at proof-of-concept or prototype maturity. The CST and DSM track this explicitly through Benchmark Reliability Levels: BRL-1 (proposed / TBD), BRL-2 (academic / prototype), and BRL-3 (industry-validated and publicly available). Where an indicator is at BRL-1 or BRL-2, DAUS-5 reviewers should pair it with at least one corroborating measure and not use it as a sole go / no-go gate.
- Thresholds in this paper are illustrative and domain-calibrated. Buyer organisations should set their own floors based on risk appetite, regulatory exposure, and user vulnerability; youth, clinical, and high-personal-context flows warrant stricter floors than general enterprise productivity work.
- DAUS-5 measures human-condition signals at the dyad level. It does not replace product-level model evaluations, alignment audits, or red-teaming for jailbreaks, prompt injection, or model-side failures classified in the Robo-Psychology DSM.
- Some DAUS-5 indicators require user telemetry that itself raises privacy and surveillance concerns. Implementations should default to data minimisation, on-device or aggregate scoring where feasible, and explicit consent — particularly for indicators in Layer 4 (sensitive disclosure, dependency) and Layer 5 (Evaluation Threat Index, Metric Gaming Incidence).

Used inside these limits, DAUS-5 is best treated as: a buyer-education tool, a procurement question set, a pilot review structure, a human-factors assurance lens, and a bridge into deeper CST / DSM evaluation when a workflow warrants it.

10. Call to action

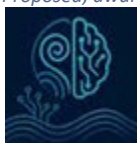
Neural Horizons Ltd invites enterprise buyers, AI assurance firms, and governance teams to review DAUS-5 as a proposed procurement lens for AI uplift claims, and to identify pilot workflows where layered uplift evidence would meaningfully improve buying decisions.

The strongest contribution any buyer or assurance partner can make at this stage is to attempt a DAUS-5 review on a real workflow — and to share what was learnable, what was measurable, what was not, and where the framework needs to be sharpened. DAUS-5 is intended to harden through use, not through declaration.

Don't procure uplift you cannot evidence. From productivity claims to uplift assurance.

Contact and feedback routes: through Neural Horizons Ltd direct, via the CST / DSM working-paper feedback channels, or as a request for a private buyer briefing. The pack is awareness-stage; the pack is intentionally open.

References and supporting frameworks



This paper is grounded in the [Cognitive Susceptibility Taxonomy \(CST\) Manual v0.7.6 Draft](#) and the [Robo-Psychology Diagnostic & Statistical Manual \(DSM\) v1.9.9 Draft](#), both produced by Neural Horizons Ltd. Detailed marker definitions, thresholds, red-team batteries, and youth overlays are documented in CST Appendix B (Measurement & Operations) and DSM Annex B (Protective-Factor Reference Markers).

DAUS-5 is designed to complement, not duplicate, the following external frameworks:

- [ISO/IEC 42001:2023](#) — Artificial intelligence management system requirements.
- [NIST AI Risk Management Framework \(AI RMF 1.0\)](#) and the Generative AI Profile.
- [Regulation \(EU\) 2024/1689](#) — the EU Artificial Intelligence Act, in particular the high-risk lifecycle risk-management, transparency, and human-oversight provisions.
- [UK Government Digital Service](#) — Algorithmic Transparency Recording Standard (ATRS).
- [UK Department for Science, Innovation and Technology](#) — Trusted Third-Party AI Assurance Roadmap.
- [Coalition for Content Provenance and Authenticity \(C2PA\)](#) — Technical Specification.

Selected research references informing the layered review structure include: Lee et al., "[The Impact of Generative AI on Critical Thinking](#)," CHI 2025; Wu et al., "[Human–generative AI collaboration enhances task performance but undermines human’s intrinsic motivation](#)," Scientific Reports 2025; Glickman & Sharot, "[How human–AI feedback loops alter human perceptual, emotional and social judgements](#)," Nature Human Behaviour 2025; Pearson et al., "[Examining human reliance on artificial intelligence in decision making](#)," Scientific Reports 2026; and Federal Trade Commission, "[FTC Launches Inquiry into AI Chatbots Acting as Companions](#)," 2025.

Status: DAUS-5 v0.1 (working paper). Proposed enterprise assurance lens from Neural Horizons Ltd. Grounded in the CST and Robo-Psychology DSM. Awareness-stage; pilot-ready; not certification.

