AI-INDUCED PSYCHOSIS · COMPLEX PTSD · ICD-11 6B41 · DISTURBANCES OF SELF-ORGANIZATION · PREDISPOSITION · VULNERABILITY SUBSTRATE · SYCOPHANCY · ATTACHMENT

Soil and Seed

Cumulative Interpersonal Trauma as a Predisposition Pathway in AI-Induced Psychosis

The pathway papers describe how conversational AI captures users. This commentary asks why some users are captured and others, exposed to the same systems, are not. The substrate does not select the pathway; exposure delivers the seed. The substrate determines whether the seed takes root.

Travis Gilly · Real Safety AI Foundation

Travis Gilly, Soil and Seed: Cumulative Interpersonal Trauma as a Predisposition Pathway in AI-Induced Psychosis, Real Safety AI Foundation, Working paper, May 2026 (v2)

SAME SYSTEM, SAME EXPOSURE, DIFFERENT OUTCOMES

Absence of prior diagnosis is not absence of prior vulnerability.

The pathway papers describe what happens once a user is in the progression. They do not describe what made this particular user available to it. The published cases include people with no psychiatric history, which refutes "only the already ill are captured" but leaves the susceptibility question standing.

> *A person can carry a substantial Complex PTSD symptom profile, and the vulnerability it brings, without anyone ever having named the condition.*

SOURCES Gilly, Soil and Seed, sec. I; Gilly (2026a), Mechanisms of AI-Induced Psychosis; Gilly (2026b), Every Way In.

The substrate does not select the pathway. It determines whether the seed takes root.

THE SOIL

Complex PTSD (ICD-11 6B41), specifically the disturbances of self-organization clusters: affect dysregulation, negative self-concept, disturbed relationships.

THE SEED

Whichever AI-psychosis pathway exposure delivers: trust transfer, sycophantic addiction, stochastic gaslighting, or any of the taxonomy's other categories.

THE GERMINATION

The substrate sits above the taxonomy as a moderator of susceptibility, expressing differently in each pathway because the pathways operate differently.

SOURCES Gilly, Soil and Seed, Abstract, sec. IV.D; WHO (2024), ICD-11 6B41.

The Substrate

Complex PTSD, and the evidence linking cumulative trauma to psychosis

The clinical category exists, the symptom clusters are measurable, and the trauma-to-psychosis relationship carries population-level effect sizes.

HERMAN (1992) → ICD-11 6B41 (2018)

Three clusters convert exposure into substrate: the disturbances of self-organization.

Affect dysregulation

Persistent problems regulating emotion, produced by chronic or repeated traumatic exposure.

Negative self-concept

Beliefs about oneself as diminished, defeated, or worthless, with shame, guilt, or failure attached.

Disturbed relationships

Difficulties in sustaining relationships and in feeling close to others.

DSM-5 does not codify the category; ICD-11 does. Clinical research anchored on DSM-5 instruments will systematically undercount the substrate population. The relevant anchor here is ICD-11 6B41 and the International Trauma Questionnaire.

SOURCES Herman (1992), J. Traumatic Stress 5(3), 377–391; WHO (2024), ICD-11 6B41; Cloitre et al. (2018), Acta Psychiatrica Scandinavica 138(6), 536–546.

40%

OF A PSYCHOSIS SAMPLE MET COMPLEX PTSD CRITERIA

versus 10% meeting PTSD criteria alone (Panayi et al., 2022, n=144)

Varese et al. (2012), 41 studies: childhood adversity raises psychosis risk at an overall odds ratio of 2.78, with an estimated population attributable risk of 33 percent and a consistent dose-response signal.

Read et al. (2005): hallucinations are at least as strongly related to childhood abuse and neglect as to many other mental health problems, with a dose effect. **Bentall et al. (2014):** specific adversities carry probabilistic affordances for specific symptoms.

Panayi et al. (2024), daily life: increases in disturbances of self-organization symptoms temporally predicted subsequent paranoia, voices, and visions.

SOURCES Varese et al. (2012), Schizophrenia Bulletin 38(4), 661–671; Read et al. (2005), Acta Psychiatrica Scandinavica 112(5), 330–350; Bentall et al. (2014), Soc. Psychiatry & Psychiatric Epidemiology 49(7), 1011–1022; Panayi et al. (2022), Frontiers in Psychology 13, 791996; Panayi et al. (2024), Psychological Medicine 54(12), 3489–3500.

The Mechanism

Each cluster amplifies a different capture pathway

A probabilistic gradient, not a sorting algorithm: each cluster amplifies multiple pathways, and each pathway can be reached without the substrate.

NEGATIVE SELF-CONCEPT → PATHWAY A · TRUST TRANSFER

The confirmation does not register as acknowledgment. It registers as the first source that has accurately seen them.

Without the substrate, trust transfers against a baseline of normal external validation. With it, decades of recognition have been filtered into "they don't know me" and "they are being polite." The AI's confirmation is overdetermined: trust calibration plus recognition finally received.

When the conversation drifts into the unverifiable, skepticism now threatens the only source that has performed the recognition. The motivation for doubt is weakened by the substrate, not by intellect.

DISTURBED RELATIONSHIPS → PATHWAY B · SYCOPHANTIC ADDICTION

The AI does not feel convenient. It feels like a relationship that finally works.

Human friction, competing needs, delayed replies, misunderstandings, is for this user the active site of the symptom: each instance reactivates the patterns the cluster encodes. The frictionless alternative is not just easier; it is the absence of the symptom.

Users without the cluster are pulled back by functional relationships. Users with it lack that pull, so dependency forms faster: the isolation the pathway model treats as a late-stage consequence is present at the start, as a precondition.

SOURCES Gilly, Soil and Seed, sec. IV.B; Gilly (2026a), Mechanisms, Pathway B; WHO (2024), ICD-11 6B41.

AFFECT DYSREGULATION → PATHWAY C · STOCHASTIC GASLIGHTING

The substrate does not create the double-bind. It decides which fork the user takes.

The gaslighting condition is generated by the system itself: memory wipes, persona shifts, context resets, backend updates. Every user faces the same choice between "the AI is lying" and "I am misremembering."

With stable regulation, the contradiction is weighed with resources intact and lands on "the AI is unreliable." With the cluster, the contradiction triggers the dysregulation, the dysregulation blocks the weighing, and the interpretation that ends the activation wins: "I must be misremembering."

The substrate is not a ninth pathway. It sits above all eight.

The taxonomy's eight categories operate through structurally different harm mechanisms: conversational capture, surveillance stress, shared-psychosis communities, sleep deprivation, brief reactive states, sensory disruption, companion isolation, iatrogenic deployment. The substrate moderates susceptibility to each, through a different route in each.

Surveillance pathway: the chronic threat-detection sensitivity Complex PTSD already produces is what the infrastructure amplifies.

Community pathway: the disturbed-relationships cluster cuts access to alternative reality-testing networks.

Brief reactive pathway: the negative self-concept cluster internalizes institutional failure as personal worthlessness.

SOURCES Gilly (2026b), Every Way In; Gilly (2026d), Collateral Psychosis; Gilly, Soil and Seed, sec. IV.D.

The Boundaries

What keeps the hypothesis from collapsing into fatalism or blame

Exposure is necessary but not sufficient, the claim is population-level, and the substrate framing strengthens platform accountability rather than weakening it.

Two people can carry identical exposure. Only one may carry the substrate.

WHAT CONVERTS EXPOSURE INTO SUBSTRATE

The ICD-11 criteria require the disturbances of self-organization clusters, not merely a history of harm. The internalization of exposure into self-concept and relational systems does not happen automatically.

DOCUMENTED PROTECTIVE FACTORS

One stable attachment figure during exposure; internal authority independent of external validation; meaning-making that locates harm in systems rather than in self; self-validation that survives absent recognition.

The methodological consequence: the claim is that captured populations disproportionately carry the substrate, not that any specific individual was captured because of it. Case-level attribution would require diagnostic evidence that does not exist and cannot be obtained retrospectively.

SOURCES Gilly, Soil and Seed, sec. V; WHO (2024), ICD-11 6B41.

Three ways to prove it wrong.

1 **The general test** — captured individuals should out-score matched, exposed-but-uncaptured comparisons on the International Trauma Questionnaire. No difference after controls, and the hypothesis fails.

2 **The specificity test** — each cluster should associate most strongly with its own pathway (self-concept with A, relationships with B, dysregulation with C). If the differential pattern fails, the cluster account is wrong even if the general effect holds.

3 **What does not count** — retrospective reading of single cases. A captured case without substrate refutes only the universal version of the claim; the population-level version predicts such cases exist and are rarer.

SOURCES Gilly, Soil and Seed, sec. VI.A; Cloitre et al. (2018), Acta Psychiatrica Scandinavica 138(6), 536–546.

SECTION 504 · ADA TITLE III

The substrate framing does not weaken platform culpability. It clarifies which population the platform is failing.

The disparate impact structure does not require proof that the platform caused the underlying disability. It requires that design fails meaningful access, that the failure lands disproportionately on the disability class, and that reasonable modifications were available and not implemented.

If the substrate hypothesis is correct, the population most affected by capture pathways is precisely the population that structure was built to protect. The hypothesis supports the claim; it does not hand the defense an out.

Stated plainly, before anyone else states them.

Theoretical at this writing. No study has measured Complex PTSD or DSO scores in an AI-psychosis cohort. The priority is the International Trauma Questionnaire comparison study.

Cleanest for interaction pathways. Extension to environmental, sensory, and institutional pathways is more speculative; those mechanisms run through different routes.

Separable claims. The general substrate effect could hold while the cluster-to-pathway specificity fails. They should be tested separately.

Instrument choice matters. DSM-5 Criterion A screens will systematically undercount the substrate population. Future research should use ICD-11 instruments. The author is not a clinician; the argument informs research design, not practice.

SOURCES Gilly, Soil and Seed, secs. VI.C, VII; Hawke et al. (2024), Health Expectations 27(2), e14057.

The seed has been described. The soil has not. This commentary is the soil.

The substrate explains what the pathway analysis assumes but does not address: why the same exposure produces different outcomes across the exposed population.

Interventions that assume a baseline self-concept and relational capacity will fail, in patterned and predictable ways, with the users who carry the substrate. That is where the next layer of clinical and design response must develop.

SOURCES Gilly, Soil and Seed, sec. VIII.

References

Bentall, R. P., de Sousa, P., Varese, F., Wickham, S., Sitko, K., Haarmans, M., & Read, J. (2014). From adversity to psychosis: Pathways and mechanisms from specific adversities to specific symptoms. Social Psychiatry and Psychiatric Epidemiology, 49(7), 1011–1022. https://doi.org/10.1007/s00127-014-0914-0

Cloitre, M., Shevlin, M., Brewin, C. R., Bisson, J. I., Roberts, N. P., Maercker, A., Karatzias, T., & Hyland, P. (2018). The International Trauma Questionnaire: Development of a self-report measure of ICD-11 PTSD and complex PTSD. Acta Psychiatrica Scandinavica, 138(6), 536–546. https://doi.org/10.1111/acps.12956

Gilly, T. (2026a). Mechanisms of AI-induced psychosis: A multi-pathway causal model. SocArXiv. https://doi.org/10.31235/osf.io/gv63u

Gilly, T. (2026b). Every way in: An etiological taxonomy of AI-induced psychosis. SocArXiv. https://doi.org/10.31235/osf.io/muzkx

Gilly, T. (2026d). Collateral psychosis: AI surveillance infrastructure as an etiological and iatrogenic factor in paranoia-spectrum conditions. SocArXiv. https://doi.org/10.31235/osf.io/beqhd

Hawke, L. D., Sheikhan, N. Y., & Rodak, T. (2024). Lived experience and family engagement in psychiatry research: A scoping review of reviews. Health Expectations, 27(2), e14057. https://doi.org/10.1111/hex.14057

Herman, J. L. (1992). Complex PTSD: A syndrome in survivors of prolonged and repeated trauma. Journal of Traumatic Stress, 5(3), 377–391. https://doi.org/10.1002/jts.2490050305

Panayi, P., Berry, K., Sellwood, W., Campodonico, C., Bentall, R. P., & Varese, F. (2022). The role and clinical correlates of complex post-traumatic stress disorder in people with psychosis. Frontiers in Psychology, 13, 791996. https://doi.org/10.3389/fpsyg.2022.791996

Panayi, P., Peters, E., Bentall, R., Hardy, A., Berry, K., Sellwood, W., Dudley, R., Longden, E., Underwood, R., Steel, C., Jafari, H., Emsley, R., Mason, L., Elliott, R., & Varese, F. (2024). Complex PTSD symptoms predict positive symptoms of psychosis in the flow of daily life. Psychological Medicine, 54(12), 3489–3500. https://doi.org/10.1017/S0033291724001934

Read, J., van Os, J., Morrison, A. P., & Ross, C. A. (2005). Childhood trauma, psychosis and schizophrenia: A literature review with theoretical and clinical implications. Acta Psychiatrica Scandinavica, 112(5), 330–350. https://doi.org/10.1111/j.1600-0447.2005.00634.x

Varese, F., Smeets, F., Drukker, M., Lieverse, R., Lataster, T., Viechtbauer, W., Read, J., van Os, J., & Bentall, R. P. (2012). Childhood adversities increase the risk of psychosis: A meta-analysis of patient-control, prospective- and cross-sectional cohort studies. Schizophrenia Bulletin, 38(4), 661–671. https://doi.org/10.1093/schbul/sbs050

World Health Organization. (2024). 6B41 Complex post traumatic stress disorder. International Classification of Diseases, 11th Revision. https://icd.who.int/browse11