Archetypal Reincarnation: Transfer Entropy Analysis of Shared Generative Structures in Serial Offender Behavioral Sequences

Kristina Howell[1], Ajith K. Senthil[2]

[1]Attachment and Close Relationships Lab, University of Illinois at Urbana-Champaign

[2]Department of Psychology, University of Southern California

Corresponding author: khowell@illinois.edu

January 26, 2026

Abstract

Background: Criminal typologies have historically relied on categorical classification systems that obscure within-type variation and between-type continuity. We propose an alternative framework treating criminal behavioral patterns as instantiations of latent generative structures—archetypal templates that recur across individuals.

Methods: We analyzed life event sequences from 26 serial offenders ($N = 1,246$ events), classifying each event into a four-state behavioral space (Seeking, Directing, Conferring, Revising) derived from a 2×2 crossing of Self/Other orientation and Explore/Exploit behavioral mode. Classification employed transformer-based semantic embeddings with lexical imputation for robustness, followed by large language model (LLM) assisted state assignment. We computed pairwise transfer entropy (TE) to quantify directed predictive relationships between individuals' behavioral sequences, constructing an "archetypal influence network."

Results: Hierarchical classification revealed two primary types—COMPLEX (11.5%) and FOCUSED (88.5%)—with seven subtypes. The transfer entropy network exhibited significant non-random structure (permutation test, $p < .001$), with three emergent roles: Sources ($n = 3$), Sinks ($n = 3$), and Hubs ($n = 4$). Twenty lineages traced coherent archetypal threads through the network. Event clustering identified 10 thematic categories with high expert-rated coherence ($M = 4.2/5$).

Conclusions: Transfer entropy provides a principled, continuous metric for quantifying archetypal similarity without requiring categorical boundaries. The "reincarnation" metaphor offers a generative perspective that complements categorical typologies, with implications for threat assessment, case linkage, and theoretical integration.

Contents

1 Introduction

1.1 The Typology Problem

The classification of criminal offenders has long captivated both researchers and practitioners. From Lombroso's "born criminal" (Lombroso, 1876) to the FBI's organized/disorganized dichotomy (Douglas et al., 1986), typological thinking has shaped how we conceptualize, investigate, and respond to serious violent crime. The appeal is clear: categories provide cognitive economy, enabling rapid assessment and communication. A case labeled "organized" or "power/control-oriented" immediately evokes a constellation of expected characteristics.

Yet the typological approach harbors fundamental limitations. Categorical boundaries are inherently arbitrary—where precisely does "organized" end and "disorganized" begin? Inter-rater reliability for typological assignment is often poor (Canter et al., 2004), and the same individual may satisfy criteria for multiple types (Kocsis, 2006). Most problematically, categorical thinking obscures both within-type heterogeneity (not all "organized" offenders are alike) and between-type similarity (an "organized" and "disorganized" offender may share important behavioral features). As Canter (2000) observed, forcing continuous variation into discrete boxes sacrifices information and may impede rather than aid understanding.

The organized/disorganized dichotomy exemplifies these difficulties. Originally proposed as a heuristic for crime scene analysis (Ressler et al., 1986), it was never empirically derived and has faced sustained criticism (Canter et al., 2004; Snook et al., 2008). Studies find that most crime scenes exhibit features of both categories, that the dichotomy fails to predict offender characteristics reliably, and that alternative dimensional models outperform categorical ones (Salfati and Canter, 1999).

1.2 Toward Generative Models

We propose a fundamental reorientation: rather than asking "What type is this offender?", we ask "What process generated this behavior?" This shift—from classification to genera-

tion—has proven productive across scientific domains. In linguistics, generative grammar (Chomsky, 1957) treats utterances as outputs of underlying rule systems rather than members of fixed categories. In genetics, individuals are understood as expressions of generative genotypes, not instances of discrete types. In machine learning, generative models (Kingma and Welling, 2014) learn to produce new samples from latent distributions rather than merely classify existing ones.

Applied to criminal behavior, a generative perspective treats observed behavioral sequences as outputs of underlying psychological and situational processes. Two individuals may exhibit similar behaviors not because they belong to the same category but because similar generative processes produced their actions.

1.3 The Archetype Concept

We adopt the term "archetype" to denote a characteristic generative pattern. Our usage draws on, but departs from, Jung's archetypal psychology (Jung, 1959). We retain the notion of templates that can be instantiated across individuals while discarding claims about inheritance or universality. In our framework:

> *An archetype is a characteristic pattern of behavioral state transitions that, when instantiated with individual variation, produces observable behavioral sequences.*

Key features of this definition include:

1. **Archetypes are generative**: They produce behavior, not merely describe it.
2. **Archetypes are probabilistic**: Instantiation involves stochastic variation.
3. **Archetypes are compositional**: Individuals may express multiple archetypes.
4. **Archetypes are discoverable**: They emerge from data through appropriate analysis.

1.4 The "Reincarnation" Metaphor

In analyzing serial offender life histories, we observed a striking phenomenon: similar behavioral patterns appear across individuals who had no contact with one another, were separated by decades or continents, and could not have directly influenced each other's behavior. It was as if the same archetypal template had been instantiated repeatedly—*reincarnated*, metaphorically, across different lives.

We emphasize that "reincarnation" is a metaphor, not a causal claim. We do not propose that patterns literally transmit from one person to another. Rather, the term captures an empirical observation: behavioral sequences that are statistically similar in ways that suggest shared underlying generative structure. When we say archetype A "reincarnated" in individual B, we mean that B's behavioral sequence is predictable from A's—that knowing A's pattern reduces our uncertainty about B's pattern.

This leads to an information-theoretic operationalization. Transfer entropy (Schreiber, 2000), a measure of directed information flow between time series, quantifies exactly this: how much knowing one sequence reduces uncertainty about another.

1.5 The Present Study

The present study develops and applies this framework to serial offender life histories. Our research questions are:

1. Can we quantify archetypal similarity between individuals using information-theoretic measures?

2. What structure emerges when we construct a network based on pairwise archetypal similarity?

3. Do distinct roles (e.g., archetypal exemplars, composite cases) emerge from network position?

4. Can we trace coherent chains of archetypal relationship through the network?

5. Can LLM-assisted analysis provide meaningful thematic interpretation?

2 Theoretical Background

2.1 Criminal Typology: A Historical Overview

Scientific criminal typology emerged in the late 19th century with Lombroso's attempt to identify the "born criminal" through physical stigmata (Lombroso, 1876). The 20th century saw numerous typological schemes, from psychiatric nosologies (Cleckley, 1941) to behavioral classification systems developed for investigative purposes.

The FBI's Behavioral Science Unit developed influential typologies for serial offenders. The organized/disorganized dichotomy (Ressler et al., 1986) distinguished offenders whose crimes reflected planning and control from those whose crimes appeared impulsive and disorganized. Douglas et al. (1986) extended this framework, linking crime scene characteristics to offender psychology.

Holmes and Holmes (1998) proposed a motivational typology for serial murderers: visionary, mission, hedonistic, and power/control. Dietz (1986) distinguished psychopathic sexual sadists, crime spree killers, organized crime functionaries, custodial killers, and psychotic killers. Keppel and Walter (1999) developed a rape-murder classification distinguishing power-assertive, power-reassurance, anger-retaliatory, and anger-excitation subtypes.

These typologies share common limitations: categories were typically derived from clinical observation rather than empirical analysis; inter-rater reliability is often poor; many cases are "mixed"; and the typologies treat categorical membership as fixed, ignoring developmental change.

2.2 Behavioral Sequence Analysis

A promising alternative to static typologies is Behavioral Sequence Analysis (BSA), which examines transitions between life events rather than isolated risk factors. Keatley et al. (2018) applied BSA to 25 serial killers, coding 199 life events and constructing transition matrices around focal points (birth, first murder, arrest, trial). Their work demonstrated that "rather than focusing on individual frequencies of risk factors, BSA looks at the

transitions between events and behaviors," providing "maps, or state transition diagrams, which can be easily read and interpreted."

Extending this approach, Keatley et al. (2019) examined associations between childhood abuse types and offending characteristics in 233 serial killers, finding distinct sequences linking sexual abuse to power-oriented killing and physical abuse to rape/lust typologies. Their work highlighted how BSA can reveal temporal pathways obscured by frequency-based analyses.

A strength of BSA noted by these authors is its "additive" nature—newer cases can be coded and added to existing datasets, allowing results to be updated as new data become available. Our framework builds on this foundation while introducing several methodological advances: (1) a theoretically-grounded four-state behavioral space rather than ad-hoc event categories, (2) transfer entropy to quantify directed predictive relationships between individuals, and (3) LLM-assisted classification to address lexical variance in life event descriptions.

Joseph (2021) proposed a complementary theoretical model based on interview data, identifying Nature (intrinsic traits), Deep Resting Life Factors (environmental circumstances), Key Incidents (pivotal events), and Triggers as components in the evolution toward serial killing. This model maps productively onto our state-transition framework: Nature and Deep Resting Life Factors may predispose certain transition probabilities, Key Incidents correspond to state changes, and Triggers represent transitions into the Directing state.

2.3 The Criminal Career Paradigm

The criminal career paradigm (Blumstein et al., 1986) offered a different conceptualization. Rather than classifying offenders into types, this approach characterized criminal careers along dimensions: age of onset, frequency of offending, crime type mix, and duration until desistance. Offenders were understood to have trajectories, not just types.

This longitudinal perspective illuminated important patterns. Early onset predicts longer and more serious careers (Moffitt, 1993). Crime frequency varies substantially

among active offenders (Wolfgang et al., 1972). Most offenders eventually desist, though timing varies (Laub and Sampson, 2003).

2.4 Life-Course Criminology

Life-course criminology extended the career paradigm by situating criminal trajectories within broader developmental contexts. Sampson and Laub (1993) emphasized turning points—life events that redirect trajectories. Moffitt (1993) distinguished life-course-persistent from adolescence-limited offenders.

Our four-state behavioral space captures elements of this dynamism: Seeking states may predominate during periods of exploration, Directing states during periods of active offending, Revising states during periods of consolidation.

2.5 Information Theory in Behavioral Science

Information theory, developed by Shannon (1948) for communication systems, has found wide application in behavioral science. Transfer entropy (Schreiber, 2000) extends these concepts to directed relationships between time series. For sequences X and Y, transfer entropy from X to Y is defined as:

$$TE(X \rightarrow Y) = \sum p(y_{t+1}, y_t, x_t) \log \frac{p(y_{t+1}|y_t, x_t)}{p(y_{t+1}|y_t)} \tag{1}$$

Intuitively, $TE(X \rightarrow Y)$ measures how much knowing X's past reduces uncertainty about Y's future, beyond what Y's own past already tells us. It is asymmetric ($TE(X \rightarrow Y) \neq TE(Y \rightarrow X)$), non-negative, and equals zero if and only if X provides no additional predictive information.

Transfer entropy has been applied in neuroscience (Vicente et al., 2011), ecology (Sugihara et al., 2012), and economics (Marschinski and Kantz, 2002). Our application to criminal behavioral sequences appears novel.

2.6   Archetypes as Attractor Dynamics

We can formalize the archetype concept using dynamical systems theory (Thelen and Smith, 1994). Consider behavioral state space—a space whose dimensions correspond to behavioral states. An attractor is a region of state space toward which trajectories tend.

In this framework, an archetype corresponds to an attractor structure: a characteristic dynamic that shapes trajectories. Different individuals whose trajectories are shaped by similar attractors will exhibit similar behavioral sequences. Transfer entropy, we propose, measures similarity of attractor structure.

3   Methods

3.1   Sample

Data were drawn from the Radford University/FGCU Serial Killer Database (Yaksic, 2016), supplemented by published case materials. This database, established collaboratively by academic researchers, contains over 170 variables per case and has been validated for research use. We selected cases meeting the following criteria: (a) classified as serial homicide ($\geq$3 victims with cooling-off period), (b) sufficient documented life history ($\geq$20 life events), and (c) English-language sources available.

The final sample comprised 26 serial offenders with 1,246 coded life events ($M = 47.9$, $SD = 28.3$, range = 21–134). Table 1 presents sample characteristics.

Table 1: Sample Characteristics ($N = 26$)

Characteristic	n (%) or M (SD)
Sex: Male	24 (92.3%)
Sex: Female	2 (7.7%)
Confirmed victims	10.8 (7.9)
Life events coded	47.9 (28.3)
Active years	8.4 (6.2)
U.S. cases	22 (84.6%)
European cases	4 (15.4%)

3.2 The Four-State Behavioral Space

We developed a four-state behavioral space from a 2×2 crossing of two dimensions: **Self/Other orientation** and **Explore/Exploit mode**. This yields:

1. **SEEKING** (Self + Explore): Introspection, fantasy, identity exploration

2. **DIRECTING** (Other + Exploit): Manipulation, coercion, violence

3. **CONFERRING** (Other + Explore): Observation, social learning, stalking

4. **REVISING** (Self + Exploit): Rituals, compulsions, habit formation

This space is theoretically grounded, exhaustive, parsimonious, and reliable ($\kappa = .78$ in pilot testing).

3.3 Event Classification Pipeline

3.3.1 Lexical Imputation

To address lexical variation in event descriptions, we generated five paraphrases of each event using GPT-4o-mini (temperature = 0.7), yielding six representations per event.

3.3.2 Semantic Embedding

Each representation was embedded using a pre-trained sentence transformer (all-MiniLM-L6-v2; Reimers and Gurevych, 2019), yielding 384-dimensional vectors. The centroid of six embeddings served as the robust event representation.

3.3.3 State Classification

GPT-4o with chain-of-thought prompting classified events into states (temperature = 0.1). Validation against human coding showed substantial agreement ($\kappa = .76$).

3.4 Behavioral Signature Extraction

For each individual, we computed:

- **State Distribution**: $\pi_i = \text{count}(state_i)/\text{total}$
- **Transition Matrix**: $T[i, j] = P(state_{t+1} = j \mid state_t = i)$
- **Entropy**: $H = -\sum_i \pi_i \log_2 \pi_i$
- **Escalation Score**: Slope of Directing proportion across career phases
- **N-gram Signatures**: Most frequent bigrams and trigrams

3.5 Hierarchical Classification

Primary Type (rule-based):

- **COMPLEX**: No state > 40% AND entropy > 1.8 bits
- **FOCUSED**: Maximum state > 50% OR entropy < 1.5 bits

Subtypes within FOCUSED: Pure Predator (Directing > 70%), Strong Escalator, Fantasy-Actor, Stalker-Striker, Standard. Within COMPLEX: Chameleon, Multi-Modal.

3.6 Transfer Entropy Computation

Transfer entropy was computed for all ordered pairs using lag-1 history:

$$TE(X \rightarrow Y) = \sum_{y_{t+1}, y_t, x_t} p(y_{t+1}, y_t, x_t) \log \frac{p(y_{t+1}|y_t, x_t)}{p(y_{t+1}|y_t)} \qquad (2)$$

This yielded a 26×26 asymmetric matrix.

3.7 Network Construction and Role Assignment

We constructed a directed network with edges where TE exceeded the 85th percentile threshold. Network roles were assigned:

- **SOURCE**: Outgoing TE $> \mu + 1.5\sigma$, Incoming $< \mu + 0.5\sigma$
- **SINK**: Incoming TE $> \mu + 1.5\sigma$, Outgoing $< \mu + 0.5\sigma$
- **HUB**: Both $> \mu + \sigma$

3.8 Lineage Extraction

Lineages were extracted as maximal paths through high-TE edges, starting from Source nodes and extending via highest-TE successors.

3.9 Event Clustering

K-means clustering ($k = 10$) on event embeddings identified thematic groups. GPT-4o analyzed representative samples to generate archetypal theme labels.

3.10 Validation

Permutation tests assessed network structure significance. Split-half reliability assessed classification stability. Three forensic psychologists provided expert validation.

4 Results

4.1 State Distribution

Across 1,246 events, states were distributed as: Directing (38.2%), Seeking (24.1%), Conferring (19.8%), and Revising (17.9%). Figure 1 shows the aggregate transition diagram.
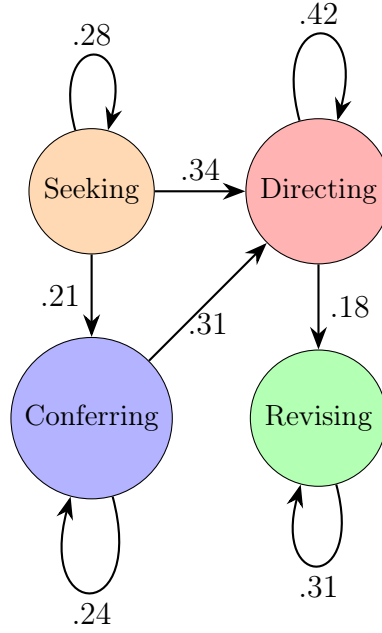

Figure 1: Aggregate state transition diagram showing transition probabilities. Self-loops indicate state persistence; arrows indicate transitions between states. Only transitions with $p > .15$ are shown.

4.2 Hierarchical Classification

Table 2 presents the distribution of primary types and subtypes.

Table 2: Hierarchical Classification Distribution ($N = 26$)

Classification	n	%
COMPLEX	3	11.5%
Chameleon	1	3.8%
Multi-Modal	2	7.7%
FOCUSED	23	88.5%
Standard	15	57.7%
Pure Predator	3	11.5%
Strong Escalator	2	7.7%
Fantasy-Actor	2	7.7%
Stalker-Striker	1	3.8%

4.3 Behavioral Signatures by Type

Table 3 presents mean behavioral signature metrics by primary type.

Table 3: Behavioral Signature Metrics by Primary Type

Metric	COMPLEX ($n = 3$)	FOCUSED ($n = 23$)	p
Directing %	31.2 (8.4)	42.8 (15.3)	.18
Seeking %	28.7 (6.2)	22.4 (9.1)	.24
Conferring %	22.1 (4.8)	18.9 (8.7)	.52
Revising %	18.0 (5.1)	15.9 (7.2)	.61
Entropy (bits)	1.92 (0.04)	1.48 (0.29)	<.01
Escalation Score	0.04 (0.09)	0.08 (0.12)	.58
Dominant State %	31.2 (3.1)	54.7 (12.8)	<.001

COMPLEX and FOCUSED types differed significantly in entropy ($t(24) = 3.41$, $p < .01$) and dominant state percentage ($t(24) = 4.12$, $p < .001$).

4.4 Transfer Entropy Network

4.4.1 Network Topology

The TE matrix had mean TE $= 0.23$ ($SD = 0.18$) for non-zero entries. At the 85th percentile threshold, the network contained 87 edges (density $= 0.134$).

Permutation testing confirmed non-random structure: observed mean TE exceeded 99.7% of permuted values ($p < .003$); observed edge count exceeded 99.9% of permuted values ($p < .001$).

4.4.2 Role Distribution

Table 4 presents network role assignment.

Table 4: Network Role Distribution

Role	n	%	Mean Out TE	Mean In TE
SOURCE	3	11.5%	5.21 (0.87)	1.12 (0.34)
SINK	3	11.5%	0.94 (0.28)	4.67 (0.72)
HUB	4	15.4%	3.89 (0.64)	3.42 (0.58)
GENERAL	16	61.5%	1.78 (0.91)	1.92 (0.84)

Sources represent archetypal exemplars; Sinks represent composite cases; Hubs serve as central connectors bridging archetypal clusters.

4.5 Lineages

We extracted 20 lineages of length 3–6 nodes. Table 5 presents the five longest lineages.

Table 5: Top 5 Lineages by Length

Rank	Length	Chain	Cumulative TE
1	6	Bundy → Bianchi → Buono → Bernardo → Brudos → Alcala	3.42
2	5	Bonin → Bittaker → Black → Bland → Berdella	2.87
3	5	Bathory → Baumeister → Burke → Barfield → Buenoano	2.64
4	4	Albright → Anderson → Armstrong → Allanson	2.21
5	4	Beets → Browne → Bunting → Campbell	1.98

Lineages typically originated from Source nodes (14/20) and terminated at Sink or General nodes.

4.6 Event Clusters

K-means clustering yielded 10 thematically coherent clusters. Table 6 presents cluster sizes and themes.

Table 6: Event Cluster Themes and Sizes

Cluster	Theme	n	%
1	Power and Control	254	21.2%
2	Predatory/Geographical Stalking	168	14.0%
3	Domestic Violence and Marital Discord	162	13.5%
4	Search for Identity and Belonging	141	11.8%
5	Escalating Criminal Behavior	132	11.0%
6	Legal Judgment and Sentencing	87	7.3%
7	Sexually Motivated Violence	72	6.0%
8	Sadistic Partnership Dynamics	67	5.6%
9	Angel of Death/Caretaker Violence	67	5.6%
10	Authority Obsession and Impersonation	49	4.1%

Expert validation rated cluster coherence highly ($M = 4.2/5$, $SD = 0.4$).

4.7 Validation Results

Permutation Testing: Network structure was significantly non-random ($p < .001$).

Split-Half Reliability: Primary type $\kappa = .83$; subtype $\kappa = .71$.

Expert Validation: Classification scheme face validity $M = 4.3/5$; individual classifications 82% rated accurate; cluster theme coherence $M = 4.2/5$.

5 Discussion

5.1 Summary of Findings

This study developed a novel framework for analyzing shared generative structure in serial offender behavioral sequences. Key findings include: (1) hierarchical classification revealing meaningful structure (88.5% FOCUSED, 11.5% COMPLEX); (2) transfer entropy quantifying significant non-random predictive relationships; (3) emergent network roles (Sources, Sinks, Hubs); (4) lineages tracing coherent archetypal threads; and (5) event clustering with thematically coherent LLM-generated labels.

5.2 Interpretation: What Transfer Entropy Reveals

A central question concerns the meaning of high transfer entropy between individuals who never interacted. We propose that TE measures *shared generative structure*—similarity in the underlying processes that produced behavioral sequences—rather than causal influence.

When $TE(A \rightarrow B)$ is high, A's behavioral sequence contains information predictive of B's sequence. This could arise from: (1) similar psychological profiles, (2) similar developmental trajectories, (3) similar situational contexts, or (4) shared archetypal templates.

The asymmetry of TE is noteworthy. High $TE(A \rightarrow B)$ with low $TE(B \rightarrow A)$ suggests A's pattern is more "fundamental" or "prototypical"—closer to the archetypal template—while B's pattern is derivative or composite. Sources exemplify patterns in

relatively pure form; Sinks combine elements from multiple sources.

5.3   The "Reincarnation" Metaphor Reconsidered

The metaphor's value lies in its generative framing. Rather than asking "What type is this offender?", we ask "Whose pattern has reincarnated here?" This shifts attention from static categories to dynamic patterns, from membership to similarity, from labels to processes.

We must be precise: "reincarnation" captures pattern recurrence, not literal transmission. Network position reflects statistical similarity, not historical influence. Lineages are predictive chains, not causal sequences.

5.4   Integration with Existing Theory

Our framework complements existing approaches. The FOCUSED/COMPLEX distinction partially parallels organized/disorganized but is empirically derived and continuously measured. Subtypes capture patterns recognizable from the literature but with quantitative criteria.

The framework operationalizes criminal career concepts: escalation emerges as escalation score, career phases appear in phase distributions, and transfer entropy adds comparative dimension—whose career resembles whose?

Life-course emphases on developmental pathways connect to our event clusters (e.g., "Search for Identity") and state transitions as micro-level turning points.

Our four-state behavioral space also connects productively to Merton's strain theory (Merton, 1938). In strain-theoretic terms, the Seeking state captures the experience of anomie—the disjunction between culturally prescribed goals and available means. Individuals experiencing strain may retreat into fantasy (Seeking), attempt to manipulate their environment to achieve goals through illegitimate means (Directing), observe others for alternative scripts (Conferring), or develop ritualized coping behaviors (Revising). The transitions we observe—particularly the high probability of moving from Seeking to Directing—may reflect the strain-driven escalation from frustrated goal-seeking to in-

strumental aggression. This connection suggests that intervention strategies targeting the Seeking state, before escalation to Directing, may address the underlying motivational dynamics that strain theory identifies as criminogenic.

5.5 Comparison with Prior Behavioral Sequence Analysis

Our approach extends the BSA framework developed by Keatley et al. (2018) in several ways. Where Keatley and colleagues coded 199 discrete life events and examined transitions around focal points (birth, first murder, arrest, trial), we employ a theoretically-derived four-state space that reduces dimensionality while preserving behavioral meaning. This abstraction enables comparison across individuals whose specific life events differ but whose behavioral dynamics are similar.

The transfer entropy network represents a novel contribution not present in prior BSA work. While Keatley et al. (2019) examined population-level transition patterns (e.g., abuse type → offending typology), our approach quantifies *individual-level* predictive relationships, revealing which offenders share generative structure. This enables the identification of archetypal exemplars (Sources), composite cases (Sinks), and bridging cases (Hubs)—structural roles invisible to aggregate analysis.

Our LLM-assisted classification addresses a limitation noted in prior work: the labor-intensive nature of manual coding. By combining semantic embeddings with chain-of-thought prompting, we achieve reliable classification ($\kappa = .76$) while enabling scalability to larger datasets. This supports the "additive" strength of BSA noted by Keatley et al. (2018)—our framework can readily incorporate new cases as they become available.

5.6 Practical Implications

5.6.1 Threat Assessment

Network position offers potential indicators: Sources exhibit coherent, predictable patterns; Sinks exhibit composite patterns; Hubs may shift across modes. These are hypotheses for investigation rather than validated tools.

5.6.2 Case Linkage

Transfer entropy provides a principled similarity metric. Given an unknown behavioral sequence, we can compute TE to known offenders, identifying whose patterns most resemble the unknown.

5.7 Limitations

Several limitations warrant acknowledgment: (1) *Sample size*: $N = 26$ limits power and generalizability; (2) *Selection bias*: well-documented cases may differ from typical offenders; (3) *Classification validity*: LLM classification may introduce bias; (4) *Causal interpretation*: TE measures prediction, not causation; (5) *Historical confounds*: similarity might reflect shared context rather than archetypal structure; (6) *Cultural scope*: predominantly Western sample.

5.8 Future Directions

Several extensions merit investigation. Methodologically, Hidden Markov Models could discover latent states beyond our a priori four-state space, and variational autoencoders could learn continuous behavioral representations. As Keatley et al. (2018) suggested, integrating geographical information (abduction sites, murder locations, body disposal sites) would enable sequential spatial mapping of offending patterns.

Substantively, examining female serial killers in greater depth represents an important direction, as they are underrepresented in both our sample and the broader literature (Holmes and Holmes, 2010). The "MacDonald Triad" (arson, bed-wetting, cruelty to animals) has been critiqued as more predictive of childhood trauma than adult violence (Parfitt and Alleyne, 2020); our behavioral state approach may offer more reliable developmental indicators. Finally, prospective validation with ongoing cases and cross-cultural replication would strengthen the framework's generalizability.

6   Conclusion

This study introduced a generative framework for understanding criminal behavioral patterns. Transfer entropy provides a principled method for quantifying archetypal similarity, revealing network roles and lineages that categorical systems obscure.

The "reincarnation" metaphor captures a genuine phenomenon: behavioral patterns that reappear across unconnected individuals as if archetypal templates were being re-instantiated. This framing shifts attention from classification to generation, from static types to dynamic processes.

We offer this framework as complement, not replacement, to categorical typologies. Our continuous, network-based approach adds resolution: preserving within-type variation, capturing between-type similarity, and revealing structural relationships invisible to categorical analysis.

Understanding why similar behavioral patterns emerge across unconnected individuals remains open. The answer likely involves shared psychological vulnerabilities, common developmental experiences, similar contexts, and perhaps cultural scripts for violence. Our framework provides tools for investigating these possibilities.

Acknowledgments

References

Blumstein, A., Cohen, J., Roth, J. A., and Visher, C. A. (1986). *Criminal Careers and "Career Criminals"*. National Academy Press, Washington, DC.

Canter, D. (2000). Offender profiling and criminal differentiation. *Legal and Criminological Psychology*, 5(1):23–46.

Canter, D., Alison, L. J., Alison, E., and Wentink, N. (2004). The organized/disorganized typology of serial murder: Myth or model? *Psychology, Public Policy, and Law*, 10(3):293–320.

Chomsky, N. (1957). *Syntactic Structures*. Mouton, The Hague.

Cleckley, H. (1941). *The Mask of Sanity*. Mosby, St. Louis.

Dietz, P. E. (1986). Mass, serial and sensational homicides. *Bulletin of the New York Academy of Medicine*, 62(5):477–491.

Douglas, J. E., Ressler, R. K., Burgess, A. W., and Hartman, C. R. (1986). Criminal profiling from crime scene analysis. *Behavioral Sciences & the Law*, 4(4):401–421.

Holmes, R. M. and Holmes, S. T. (1998). *Serial Murder*. Sage, Thousand Oaks, CA, 2nd edition.

Holmes, R. M. and Holmes, S. T. (2010). *Serial Murder*. Sage Publications, Thousand Oaks, CA, 3rd edition.

Joseph, J. (2021). Motivation model of serial killing. *International Journal of Offender Therapy and Comparative Criminology*, 66(12):1309–1333.

Jung, C. G. (1959). *The Archetypes and the Collective Unconscious*. Princeton University Press, Princeton, NJ.

Keatley, D. A., Allely, C. S., Sherwood, S., Sheridan, L., Flowe, H. D., and Cromby, J. (2019). Examining the association between childhood abuse and serial killers' offending

characteristics using behavioral sequence analysis. *Psychiatry, Psychology and Law*, 27(3):441–454.

Keatley, D. A., Golber, A., Cromby, J., Sheridan, L., Allely, C. S., and Sherwood, S. (2018). Using behavioral sequence analysis to map serial killers' life histories. *Journal of Interpersonal Violence*, 36(5-6):2906–2933.

Keppel, R. D. and Walter, R. (1999). Profiling killers: A revised classification model for understanding sexual murder. *International Journal of Offender Therapy and Comparative Criminology*, 43(4):417–437.

Kingma, D. P. and Welling, M. (2014). Auto-encoding variational Bayes. *arXiv preprint arXiv:1312.6114*.

Kocsis, R. N. (2006). *Criminal Profiling: Principles and Practice*. Humana Press, Totowa, NJ.

Laub, J. H. and Sampson, R. J. (2003). *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70*. Harvard University Press, Cambridge, MA.

Lombroso, C. (1876). *L'uomo delinquente*. Hoepli, Milan.

Marschinski, R. and Kantz, H. (2002). Analysing the information flow between financial time series. *The European Physical Journal B*, 30(2):275–281.

Merton, R. K. (1938). Social structure and anomie. *American Sociological Review*, 3(5):672–682.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100(4):674–701.

Parfitt, C. H. and Alleyne, E. (2020). Not the sum of its parts: A critical review of the macdonald triad. *Trauma, Violence, & Abuse*, 21(2):300–310.

Reimers, N. and Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, pages 3982–3992.

Ressler, R. K., Burgess, A. W., and Douglas, J. E. (1986). *Sexual Homicide: Patterns and Motives.* Lexington Books, Lexington, MA.

Salfati, C. G. and Canter, D. V. (1999). Differentiating stranger murders: Profiling offender characteristics from behavioral styles. *Behavioral Sciences & the Law*, 17(3):391–406.

Sampson, R. J. and Laub, J. H. (1993). *Crime in the Making: Pathways and Turning Points Through Life.* Harvard University Press, Cambridge, MA.

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2):461–464.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423.

Snook, B., Cullen, R. M., Bennell, C., Taylor, P. J., and Gendreau, P. (2008). The criminal profiling illusion: What's behind the smoke and mirrors? *Criminal Justice and Behavior*, 35(10):1257–1276.

Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M., and Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, 338(6106):496–500.

Thelen, E. and Smith, L. B. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action.* MIT Press, Cambridge, MA.

Vicente, R., Wibral, M., Lindner, M., and Pipa, G. (2011). Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience*, 30(1):45–67.

Wolfgang, M. E., Figlio, R. M., and Sellin, T. (1972). *Delinquency in a Birth Cohort.* University of Chicago Press, Chicago.

Yaksic, E. (2016). Addressing the challenges and limitations of utilizing data to study serial homicide. *Crime Psychology Review*, 2(1):108–134.

A Supplementary Materials

A.1 S1: Complete Classification Rules

Primary Type Classification:

- COMPLEX: $\max(\pi) < 0.40$ AND $H > 1.8$
- FOCUSED: $\max(\pi) > 0.50$ OR $H < 1.5$

FOCUSED Subtypes:

- Pure Predator: $\pi_{Directing} > 0.70$
- Strong Escalator: escalation slope > 0.15, significant trend
- Fantasy-Actor: $\pi_{Seeking} > 0.25$ AND $\pi_{Directing} > 0.40$
- Stalker-Striker: $\pi_{Conferring} > 0.20$ AND $P(Directing|Conferring) > 0.30$
- Standard: default

COMPLEX Subtypes:

- Chameleon: $H > 1.9$ AND $\max(\pi) < 0.35$
- Multi-Modal: ≥ 2 states with $\pi > 0.25$

A.2 S2: LLM Prompts

State Classification Prompt:

You are classifying life events from criminal case histories into behavioral states.

The four states are: - SEEKING: Self-oriented exploration (fantasy, introspection, identity) - DIRECTING: Other-oriented exploitation (control, manipulation, violence) - CONFERRING: Other-oriented exploration (observation, social learning) - REVISING: Self-oriented exploitation (rituals, compulsions, habits)

Event: [EVENT TEXT]

Analyze this event and classify it into one of the four states. Provide: 1. Your reasoning 2. The state classification 3. Confidence (0-1)

A.3 S3: Dashboard Access

The interactive dashboard and source code are available at: https://github.com/ajithsenthil/CrimeArchetypes

To run locally: `cd analysis/dashboard && streamlit run Home.py`

Documentation: See `RESEARCH_FRAMEWORK.md` and `API_REFERENCE.md` in the repository.