Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/EU08.20250468
Published: July 2, 2025

AI-enabled Human Behavior Dynamics in Sociotechnical Manufacturing Systems: A Multimodal Ensemble Transformer

Mulang Song and Roger J. Jiao
School of Mechanical Engineering
Georgia Institute of Technology, USA

Abstract

Industry 5.0 calls for manufacturing systems that sense human behavior with the same resolution used to monitor machines so that governance can adapt in real time without undermining operator autonomy. This paper introduces an AI-enabled multimodal ensemble Transformer that fuses shop-floor video, tool-embedded signals, PLC tags and HMI events through modality-specific encoders and a cross-modal Transformer. The framework produces continuous estimates of operator compliance and supervisory demand. These real-time estimates parameterize an evolutionary cooperation–competition (ECC) game that formalizes the strategic co-adaptation of operators and engineers. Pay-off functions are expressed in terms of measurable costs, benefits and penalties; replicator dynamics are then used to study how compliance and oversight evolve under alternative incentive structures. A connector-assembly case study demonstrates how live behavioral estimates can be fed back into supervisory policy and how the combined perception–game loop helps managers balance quality, throughput and human workload. The proposed approach provides a deployable blueprint for embedding AI-powered behavioral analytics into sociotechnical manufacturing systems, advancing the human-centric and adaptive ambitions of Industry 5.0.

Keywords
Transformer, Sensor-fusion, Evolutionary Competition Cooperation Game, Sociotechnical Systems

1. Introduction
As manufacturing advances toward Industry 5.0, the production floor is no longer a collection of isolated cyber-physical modules but an integrated sociotechnical system in which human ingenuity, machine autonomy, and data intelligence co-evolve. Operators remain essential decision-makers whose creativity and judgement drive quality and resilience; however, deviations from standard operating procedures (SOPs) continue to cause a large share of defects and accidents. Addressing this gap requires instrumentation capable of describing what operators and supervisors do on the shop floor and delivering that information to higher-level control and optimization layers in real time. Realizing that vision demands instrumentation capable of perceiving what people and machines are doing and converting those perceptions into knowledge that control, and optimization layers can act on. The present study embeds such instrumentation in an AI-enabled multimodal ensemble framework and investigates how the behavioral information it produces reshapes supervisory governance and operator compliance.

1.1 AI-Enabled Multimodal Ensemble Framework
Recent advances in computer-vision action recognition, wearable activity monitoring, and low-latency edge inference now make it possible to observe human behavior on the shop floor with millisecond-level precision. The proposed multimodal transformer combines these advances into a single perception layer designed for Industry 5.0 production lines. Early work on gesture detection focused on offline quality inspection, while ergonomic studies relied on isolated IMU packs and physiological sensors, and control engineers treated PLC logs as purely machine-centric signals. Recent breakthroughs in pose-graph convolution, transformer-based self-supervision, and on-device neural

acceleration have eliminated the technical barriers that kept these modalities apart. As a result, RGB-D cameras, audio arrays, inertial wearables, and controller event streams can now be synchronized, processed, and fused within the cycle time of a modern assembly task.

1.2 Human Behavior Dynamics in Sociotechnical Systems

The behavioral state allows the study of sociotechnical system dynamics—the continuous interaction among technical subsystems, human operators, and organizational rules (Agote-Garrido et al. 2023). Action-segmentation networks now label each production step in real time (Al-Amin et al. 2019; Gudlin et al. 2024), and wearable models detect fatigue-induced motor changes. When these heterogeneous signals are fused by the multimodal ensemble transformer, managers obtain a detailed picture of shop-floor behaviors that can be used to address safety, productivity, and learning simultaneously.

1.3 Supervision–Compliance Synergy

Embedded in these dynamics is Supervision–Compliance Synergy (SCS)—the mutually adaptive relationship between operator compliance and supervisory governance. Compliance is not a static attribute, but a strategic behavior shaped by oversight intensity, incentives, peer norms, and perceived risk (Bhattacharya et al., 2023). Conversely, supervisors calibrate their interventions according to observed compliance patterns. The ensemble framework supplies each side with timely, high-resolution feedback, realizing the Industry 5.0 ideal of technology that augments—rather than overrides—human agency.

1.4 SCS as an Evolutionary Game

To formalize the adaptive interaction underlying SCS, the paper models operators and supervisors as populations engaged in an evolutionary cooperation–competition game. Evolutionary game theory reflects bounded rationality and trial-and-error learning more realistically than classical static games (McCoy 2021). Operators experiment with shortcuts to reduce cycle times; supervisors adjust inspection frequencies to control risk. Pay-offs in the game are parameterized by the ensemble-derived measures, creating a closed loop between real-time observation and strategy evolution.

1.5 Research Scope and Contributions

The work makes four specific contributions. First, it presents a scalable ensemble pipeline that fuses heterogeneous sensor modalities into a single behavioral descriptor suitable for control applications. Second, it integrates that descriptor with a formal definition of SCS rooted in sociotechnical theory. Third, it develops an evolutionary game model augmented with payoffs to analyze SCS stability and sustainability.

The rest of paper is organized as follows. Section 2 reviews literature on related fields. Section 3 elaborates the sociotechnical system dynamics of manufacturing systems. Section 4 introduces the architecture of the proposed multimodal ensemble Transformer. Section 5 develops the Evolutionary Cooperation–Competition (ECC) game. Section 6 introduces stability analysis. Section 7 concludes with managerial implications and future research directions aligned with Industry 5.0. By embedding a multimodal ensemble Transformer into behavioral game dynamics, the paper offers a rigorous yet practical blueprint for managing human–machine interactions in sociotechnical manufacturing, advancing the human-centric and adaptive imperatives of next-generation industry.

2. Literature Review

2.1 Game Decisions in Industry 5.0

Game theory has increasingly been applied to manufacturing systems to capture the strategic interactions among different decision-makers, such as machines, humans, or firms in production scenarios (Renna, 2024). Classic applications include production planning games, resource allocation competitions, and scheduling conflicts (Zhang et al., 2017), in which each agent pursues its own objectives, and a game-theoretic solution represents a balanced outcome (Liu et al., 2024). Notably, cyber–physical production systems (CPPS) have been identified as a natural domain for multi-agent and game-theoretic methods due to the presence of distributed intelligent components and actors (Tushar et al., 2023). CPPS are well-suited to game-theoretic methods because of their integration of IoT-enabled machines and real-time data analytics tools (Li & Whang, 2001).

However, prior work often prioritizes machine–agent interactions over human-centric dynamics, despite Industry 5.0's emphasis on human operator agency. Recent advancements recognize human operators as active agents within game-theoretic models rather than external disturbances. This shift aligns with Industry 5.0's focus on human-centric automation (Nahavandi, 2019), where operators' strategic choices—such as compliance with safety protocols or task prioritization—directly impact manufacturing performance (Gladysz et al., 2023). Previous studies have used games to analyze human–robot collaboration, demonstrating that framing human–robot task assignments as a game can improve both efficiency and safety (Tian et al., 2022). In the domain of manufacturing cybersecurity, some researchers have proposed game-theoretic decision models that incorporate human judgment in responding to cyberattacks (Baksi, 2022).

Overall, the literature suggests that Industry 5.0 scenarios benefit from game theory's ability to formalize both conflict and cooperation. However, most existing models focus either on machine–agent interactions or on high-level organizational decisions, and there is a relative lack of game-theoretic modeling of human interactions within manufacturing systems,

2.2 Sociotechnical System Dynamics

Sociotechnical system dynamics, rooted in the foundational work of the Tavistock Institute (Trist & Bamforth, 1951), emphasizes the interdependence of social and technical subsystems in shaping organizational outcomes. It investigates the co-evolution of human behavior and technological infrastructure, highlighting how decision-making, collective interaction, and technological affordances jointly shape complex adaptive systems. Early studies framed workplaces as joint systems in which human agencies and technical structures co-evolve to optimize productivity and safety (Cherns, 1976).

Recent advances emphasize population-level behavior as a key mechanism within sociotechnical system dynamics. Rather than analyzing individuals in isolation, current models focus on how groups adapt behavior through feedback loops between digital infrastructures and social influence. Xiao and Zhang (2023) model government–internet user interactions in networked mass events as an evolutionary game, where policy adjustments and public compliance iteratively shape the spread of information. Similarly, Liu and Xiao (2019) demonstrate how opinion leaders within social networks guide public opinion reversal by dynamically influencing collective adherence.

Researchers have also employed system dynamics and simulation techniques to study how human decisions feed back into manufacturing system behavior over time. Manufacturing systems are inherently complex sociotechnical systems characterized by interdependencies, time delays, and non-linear effects. Assuad et al. (2020) used learning factory simulations to show that engineering students acting as production operators could tweak production parameters in a realistic simulator and observe the resulting ripple effects on throughput, inventory, and lead times.

2.3 Compliance Monitoring and Supervision

In safety-critical domains, effective supervision and rigorous compliance measures work in tandem to uphold high operational and safety standards. The interaction between oversight mechanisms and worker adherence to protocols has been widely studied across industries. In healthcare, for instance, strong safety leadership and structured oversight have been linked to higher compliance with best practices among staff (Arboh et al., 2024). Similarly, in construction, on-site supervisors and safety managers play a critical role in enforcing rules and modeling safe behavior. Research shows that clear communication by supervisors, when supported by strong organizational safety culture, significantly increases adherence to safety protocols (Yan & Zhao, 2023).

With the onset of digital transformation, traditional approaches to supervision and compliance are being redefined in smart manufacturing contexts. Modern factories are leveraging cyber-physical systems (CPS), IoT sensors, and AI to enforce real-time compliance with safety and quality standards. For example, advanced sensor networks can continuously monitor environmental and ergonomic conditions to flag potential violations (Svertoka et al., 2021). AI-based compliance monitoring tools are increasingly adopted to observe operator behavior and adherence to standard operating procedures (Indris et al., 2025). Vision-based systems can track each production step, automatically detecting incorrect actions, safety lapses, or deviations from assembly sequences. In high-mix, manual assembly environments, studies have shown that such tools help reduce defects and ensure consistent process compliance (Chen et al., 2020).

While compliance technologies offer valuable visibility into operator behavior, their broader impact lies in enabling data-driven supervisory strategies. Compliance data can inform real-time supervisory decisions and guide workflow adjustments, contributing to a continuous feedback loop between oversight and operator behavior. Emerging literature increasingly highlights how such digitally mediated interactions between monitoring systems and human behavior reflect the evolution of supervision and compliance into an integrated sociotechnical function within smart manufacturing.

2.4 Evolutionary Game Theory

Game theory has traditionally served as a foundational tool for modeling strategic interactions among rational agents, with its classical roots in economics and decision science, where players are assumed to make optimal decisions under fixed rules and complete information. However, the assumption of static rationality limits its applicability in dynamic environments. This challenge was addressed by the emergence of evolutionary game theory (EGT), which reframes strategic decision-making as a population-level process driven by selection, imitation, and mutation. Initially applied to biological systems such as predator–prey dynamics and population genetics, EGT has since expanded to model cooperation in social dilemmas (Taylor & Nowak, 2007).

In industrial contexts, EGT has proven valuable for understanding adaptive dynamics among decentralized agents. Applications include modeling cooperation–competition trade-offs in crowdsourcing (Song, 2024) and strategy evolution in platform-based ecosystems. Ji et al. (2015) employed evolutionarily stable strategies (ESS) to analyze customer–manufacturer relationships. Despite this progress, industrial decision-making often involves ambiguous human judgments that resist crisp quantification. Operators evaluate costs, efforts, and risks based on subjective perceptions shaped by context and experience.

Recent studies have begun to apply evolutionary game models to operator compliance in safety-critical domains. Huang et al. (2022) modeled construction worker behavior versus managerial enforcement, revealing stable equilibria dependent on initial compliance levels and incentive mechanisms.

3. Sociotechnical System Dynamics of Manufacturing Systems

Production at the mid-sized electronics facility is governed by a continuous trade-off between operator discipline and engineering oversight. Operators, working under tight takt times, sometimes view the compliance cost of fully following standard operating procedures as a direct impediment to throughput and are therefore tempted to omit in-process checks. Engineers confront their own supervision cost: dedicating additional labor to inspection improves quality assurance but diverts resources from maintenance and process optimization. Because the payoff for either side depends on the other's current behaviors, the shop floor functions as a strategic sociotechnical system. When most operators adhere to SOPs, the marginal value of heavy monitoring falls and engineers rationally scale back inspections; when deviations become common, increased oversight becomes the economically sound response. Even small shifts in these mutually adaptive choices influence scrap, rework and warranty exposure across the entire line.

Figure 1 situates these interactions at the task-execution stage. Peer Support and Influence shapes collective operator behavior, while Task Setup and Operation captures direct operator–machine contact that determines real-time compliance. Within the engineering population, Collaboration and Coordination governs the distribution of supervisory effort, and Maintenance and System Configuration modulates machine readiness, indirectly affecting the feasibility of SOP adherence. Performance Accountability and Validation links the two human groups by using machine fault logs and task data to recalibrate future inspection intensity and responsibility attribution. These social and technical signals feed into Compliance Enforcement and Supervisory Strategy, the levers of Supervision–Compliance Synergetic Management. Their combined effectiveness is ultimately reflected in Manufacturing Performance indicators such as OEE, MTTR and MTBF, revealing whether the system is converging on stable high quality or drifting toward defect-prone volatility.

Traditional performance metrics appear only after defects have occurred, leaving limited opportunity for timely intervention. The plant therefore requires a mechanism that estimates operator compliance as tasks unfold. The following section introduces a multi-sensor ensemble framework that meets this need by fusing video feeds, tool-mounted torque signatures, PLC events and other shop-floor signals to produce a continuous compliance score. This real-time estimate closes the loop between observed behavior and supervisory policy, enabling engineers to adjust oversight dynamically and allowing the sociotechnical game model to update its payoffs.

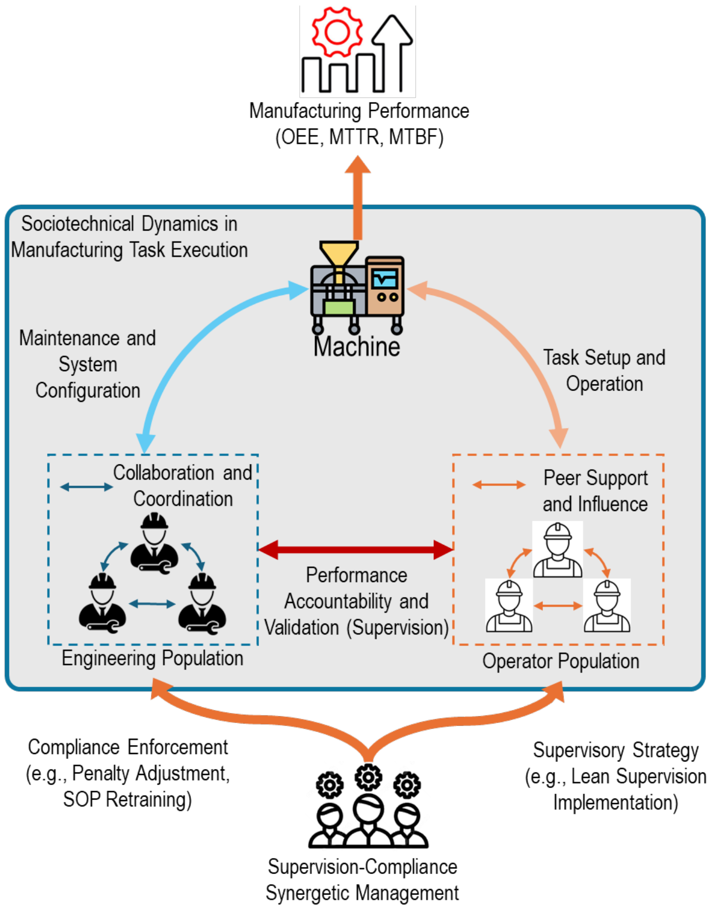

Figure 1. Sociotechnical aspects of manufacturing systems

4. A Multimodal Ensemble Transformer for Sociotechnical Systems

The architecture of the multimodal ensemble Transformer is shown in Figure 1 provides the perceptual backbone required to close the supervision–compliance loop modelled by the ECC game. It converts raw video, tool and controller signals into a pair of continuous probabilities that quantify, in real time, the proportion of operators who are following the prescribed SOP and the proportion of engineers who must intervene. These probabilities form the state variables of the replicator dynamics, allowing the theoretical game to operate on live shop-floor evidence rather than assumed constants.

At the top of the stack, the data-collection layer acquires four complementary streams: RGB-D video captures body-pose and hand-tool interactions, smart-tool inertial and torque sensors provide force-signature traces, PLC tags report machine-state transitions at millisecond granularity, and the electronic work-instruction terminal emits a time-stamped record of menu acknowledgements and checklist completions. Because each device maintains its own internal clock and sampling rate, an intermediate synchronization module aligns every packet to a common two-second sliding window. The alignment engine applies line-controller NTP stamps, resamples high-rate signals such as torque from one kilohertz to the fifty-hertz master timeline, aggregates PLC bits into forty-millisecond bins and inserts padding tokens where a modality is temporarily absent. The outcome is a set of co-registered tensors in which every row represents the same physical instant across all channels.

Once temporally aligned, each modality is passed to a specialized encoder. Video clips enter an inflated 3-D ResNet (I3D) that inflates ImageNet kernels along the temporal axis to preserve motion cues; the network outputs a 128-dimensional embedding for each frame in the window. Tool torque and current profiles are processed by a causal one-dimensional convolutional network whose dilated filters capture the rundown slope and stall signature characteristic of compliant fastening. PLC and HMI event sequences, inherently symbolic, are mapped through a gated recurrent unit that encodes step order and operator acknowledgements into a compact thirty-two-dimensional vector. These modality-specific embeddings are concatenated with a learnable task token that identifies the active machine and SOP step, forming a unified sequence that proceeds to the fusion stage.

Cross-modal integration is achieved with a four-layer Transformer encoder whose positional encoding preserves temporal order. Multi-head self-attention dynamically weights each modality according to instantaneous signal quality so that, for example, the torque embedding dominates during a tightening operation while the vision embedding dominates during an inspection gesture. Residual feed-forward blocks refine the representation without compromising inference latency (Figure 2).

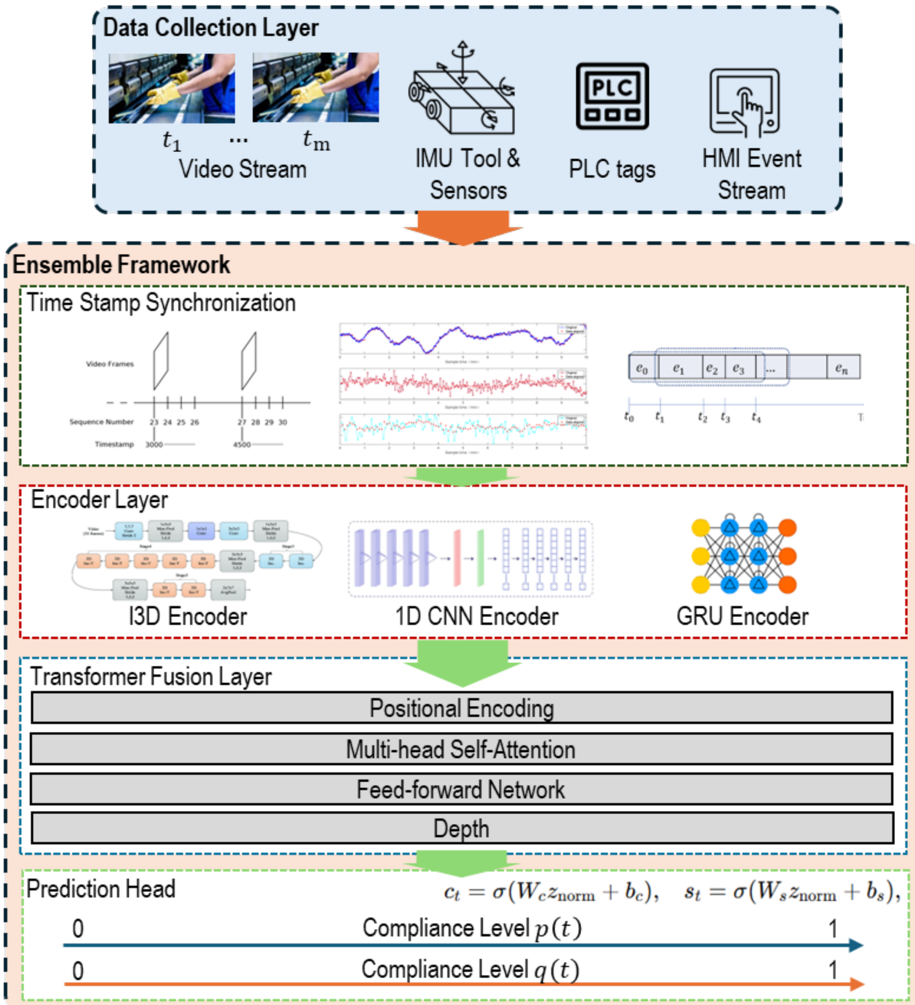

Figure 2. Process of supervision-compliance state prediction with multimodal data ensemble Transformer

Following the cross-modal Transformer, the fused sequence is headed by the special classification token $h_{CLS} \in R^{256}$, whose vector contains the context that the self-attention mechanism has distilled from every modality and every time-step in the two-second window. The prediction head converts this vector into continuous estimates of operator-side compliance and engineer-side supervision demand. First, h_{CLS} is projected into a compact bottleneck through

$$z = \mathrm{GELU}(\mathrm{W}_1 h_{CLS} + b_1), W_1 \in R^{128\times 256}, b_1 \in R^{128}$$

where the Gaussian-error linear unit supplies a smooth non-linearity that has proved superior to ReLU in handling mixed numerical scales. The 128-dimensional activation z is then stabilized by layer normalization and lightly regularized with 10 % dropout. Two parallel linear branches translate this shared representation into scalars as shown in the two equations below.

$$p_t = \sigma\left(W_p z_{norm} + b_p\right), W_c \in R^{1\times128}, \qquad q_t = \sigma\left(W_q z_{norm} + b_q\right), W_q \in R^{1\times128}$$

where $\sigma(x) = \frac{1}{1+e^{-x}}$ constrains the outputs to the unit interval. During training the head minimizes a composite objective. At inference the network emits p_t and q_t every 500 ms; these values map directly to the dynamic proportions p(t)p(t)p(t) and q(t)q(t)q(t) in the ECC replicator model, thereby coupling the learned perception stack to the analytical supervision–compliance dynamics.

5. Evolutionary Cooperation-Competition Game Model for Sociotechnical System Dynamics

This section develops and analyses an Evolutionary Cooperation–Competition (ECC) game to describe sociotechnical system dynamics in the task execution stage of the shop floor. Section 5.1 introduces the payoffs of the two-player, single-machine game: operator compliance and engineering supervision are expressed through payoff functions, and their strategic trajectories are examined with replicator dynamics. Section 5.2 introduced the replicator equations derived based on the model payoffs.

5.1 Model Payoffs of 2-player, Single Machine Game

This subsection constructs a two-player Evolutionary Cooperation–Competition (ECC) game that serves as the analytical basis for supervision–compliance synergy (SCS) on a single type of machine. Operators (Op) may comply with the standard operating procedure (SOP) or deviate for short-term throughput gains, whereas Engineers (Eng) may supervise or neglect the operation, weighing inspection resources against the risk of future rework. To support the modeling framework, the following assumptions are established:

1. The population of operators (Op) is always larger than the population of engineers (Eng);
2. The population of operators and engineers remains relatively constant.
3. Operators and engineers do not coordinate strategies.
4. Supervision can be conducted on and by any agent within the two-population system.
5. Machine reliability is static in the manufacturing environment.
6. Agents can choose between two strategies based solely on payoff comparisons.

Let $p(t) = \frac{X_1}{Op}$, $q(t) = \frac{X_2}{Eng}$ denote, respectively, the fraction of operators who comply with SOPs and the fraction of engineers who engage in active supervision at time t. The engineer-to-operator ratio $r = \frac{Eng}{Op}$, 0<r<1 together with supervision intensity $\theta \in (0,1)$ yields the effective supervision coverage $\phi = rq\theta$, $\theta \in (0,1)$ the average probability that an individual operator's behavior is inspected during a production cycle.

Adopting Friedman's state-space notation, the four strategic combinations are:

$$\mathbb{S} = \left\{\left(C_{op}, S_{Eng}\right), \left(C_{op}, N_{Eng}\right), \left(D_{op}, N_{Eng}\right), \left(D_{Op}, N_{Eng}\right)\right\}$$

Here C_{Op} and D_{Op} denote compliance and deviation, while S_{Eng} and N_{Eng} denote supervision and neglect. Operators adjust $p(t)$ by weighing the benefit of faster completion against potential penalties, and engineers adjust $q(t)$ by balancing inspection cost against the likelihood and impact of deviations. The resulting pay-off matrix as shown in Table 1 lists $\pi_{Op}(\cdot,\cdot)$ and $\pi_{Eng}(\cdot,\cdot)$ for each strategy pair, providing the foundation for replicator-dynamics analysis and subsequent stability evaluation (Table 1).

Table 1. Payoff Matrix in Two-Player ECC Game Model

Operators	Engineers	
	X_1 Engineers choose to Supervision S	$Op - X_1$ Engineers choose to Neglect N
X_2 Operators choose to comply C	$(\pi_{Op}(C,S), \pi_{Eng}(C,S))$	$(\pi_{Op}(C,N), \pi_{Eng}(C,N))$
$Eng - X_2$ Operators choose to deviate D	$(\pi_{Op}(D,S), \pi_{Eng}(D,S))$	$(\pi_{Op}(D,N), \pi_{Eng}(D,N))$

5.2 Replicator Equations of the 2-player, Single Machine Game

In the two-player, single-machine ECC setting, operator and engineer pay-offs depend on a consistent set of economic factors that govern their strategic interaction. Each operator is entitled to a base compensation Ω_{Op} for routine duties, just as the engineering cohort receives a baseline return Ω_{Eng} for fulfilling its core responsibilities. Adhering to an SOP imposes a compliance cost α_{Op} on the operator, whereas deviation yields an immediate benefit β_{op} by shortening cycle time. If that deviation is discovered, the operator is fined γ_{Op}. Engineers, when supervising, incur an effort cost α_{Eng} that scales with their chosen inspection intensity θ; when a deviation slips through, they absorb a rework cost β_{Eng}.

The operator's pay-off function is summarized in the equation below. A compliant operator earns $\Omega_{op} - \alpha_{op}$. A deviating operator gains $\beta_{op}(1-p)$—the term $(1-p)$ reflects risk dilution as deviations spread—and, if caught, pays $\gamma_{op}\phi$.

$$\pi_{Op} = \begin{cases} \Omega_{op} - \alpha_{op}\,, (C,S) \\ \Omega_{op} - \alpha_{op}, (C,N) \\ \Omega_{op} + \beta_{op}(1-p) - \gamma_{op}\phi, (D,S) \\ \Omega_{op} + \beta_{op}(1-p), (D,N) \end{cases}$$

Equation below gives the engineer's pay-off. Supervising yields $\Omega_{Eng} - \alpha_{Eng}\frac{\theta}{q}$; if a deviation is detected, a further penalty $\gamma_{Eng}\phi$ is borne. When engineers neglect supervision and operators deviate, a fault-recovery penalty $\beta_{Eng}(1-p)$ applies:

$$\pi_{Eng} = \begin{cases} \Omega_{Eng} - \alpha_{Eng}\dfrac{\theta}{q}\;\;, (C,S) \\ \Omega_{Eng}, (C,N) \\ \Omega_{Eng} - \alpha_{Eng}\dfrac{\theta}{q} - \gamma_{Eng}\phi, (D,S) \\ \Omega_{Eng} - \beta_{Eng}(1-p), (D,N) \end{cases}$$

Because an operator cannot know in advance whether supervision will occur, expected payoffs are weighted by ϕ:

$$\begin{cases} \pi_{Op}(C) = \phi\ \cdot \pi_{Op}(C,S) + (1-\phi)\pi_{Op}(C,N) \\ \pi_{Op}(D) = \phi\ \cdot \pi_{Op}(D,S) + (1-\phi)\pi_{Op}(D,N) \end{cases}$$

Engineers, in turn, condition their expected return on the fraction of deviating operators $(1-p)$:

$$\begin{cases} \pi_{Eng}(S) = p \cdot \pi_{Eng}(C,S) + (1-p)\pi_{Eng}(D,S) \\ \pi_{Eng}(N) = p\ \cdot \pi_{Eng}(C,N) + (1-p)\pi_{Eng}(D,N) \end{cases}$$

Average population pay-offs follow directly:

$$\bar{\pi}_{Op} = p\pi_{Op}(C) + (1-p)\pi_{Op}(D), \qquad \bar{\pi}_{Eng} = q\pi_{Eng}(S) + (1-q)\pi_{Eng}(N)$$

In evolutionary game theory, the change in a strategy's prevalence is governed by its fitness advantage—defined as the difference between that strategy's payoff and the average payoff of its population (Hofbauer & Sigmund 1998; Hauert et al. 2002). Accordingly, the replicator-dynamic system for the operator–engineer game is

$$\frac{dp}{dt} = p\big(\pi_{Op}(C) - \bar{\pi}_{Op}\big), \qquad \frac{dq}{dt} = q\big(\pi_{Eng}(S) - \bar{\pi}_{Eng}\big)$$

where $p(t)$and $q(t)$ are, respectively, the fractions of compliant operators and supervising engineers. Substituting the payoff expressions yields the coupled differential system

$$\frac{dp}{dt} = p(1-p)\big(-\alpha_{Op} - \beta_{Op}(1-p) + \gamma_{Op}^2\phi\big), \qquad \frac{dq}{dt} = q(1-q)\left(-\frac{\alpha_{Eng}\theta}{q} + (1-p)\big(\beta_{Eng} - \gamma_{Eng}\phi\big)\right)$$

with ϕ denoting effective supervision coverage. Setting both derivatives to zero produces up to five equilibrium points in the unit square $[[0,1] \times [0,1]$: four on the boundaries$e_1 = (0,0)$, $e_2 = (0,1)$, $e_3 = (1,1)$, $e_4 = (1,0)$, and at most one interior equilibrium p^*. The engineer coordinate q^* satisfies the cubic

$$aq^{*3} + bq^{*2} + cq^* + d = 0, \qquad \begin{cases} a = \gamma_{Op}^2\gamma_{Eng}r^2\theta^2 \\ b = -\big(\gamma_{Op}^2\beta_{Eng}r\theta + \alpha_{Op}\gamma_{\mathrm{Eng}}r\theta\big) \\ c = -\alpha_{Op}\beta_{Eng} \\ d = -\beta_{Op}\alpha_{Eng}\theta \end{cases}$$

and the corresponding operator coordinates as follows

$$p^* = 1 - \frac{-\alpha_{Op} + \gamma_{Op}^2 rq\theta}{\beta_{Op}}$$

Equations trace the evolutionary trajectories of compliance and supervision and identify the conditions under which the system stabilises, providing the baseline for subsequent population-dynamics analysis.

6. Stability Analysis of ECC Game Model

In the two-player, single-machine ECC game the only candidates for long-run outcomes are the equilibria of the replicator dynamics, yet an equilibrium is evolutionarily stable only if it withstands infinitesimal perturbations. Local stability is assessed by the Jacobian Jevaluated at an equilibrium (p^*, q^*). Linearization of System yields the matrix shown below. The trace–determinant test for planar systems (Hirsch et al., 2013) states that an equilibrium is locally asymptotically stable—and therefore an evolutionarily stable strategy (ESS) when $\det J$ and $\operatorname{tr} J$; a negative determinant identifies a saddle, while a positive determinant together with a positive trace indicates an unstable node or focus.

$$J = \begin{bmatrix} \frac{\partial \dot{p}}{\partial p} & \frac{\partial \dot{p}}{\partial q} \\ \frac{\partial \dot{q}}{\partial p} & \frac{\partial \dot{q}}{\partial q} \end{bmatrix}$$

$$\frac{\partial \dot{p}}{\partial p} = -(2p-1)\Upsilon_2 - \beta_{Op}p(p-1), \frac{\partial \dot{p}}{\partial q} = -\Upsilon_1^2 pr\theta(p-1)$$

$$\frac{\partial \dot{q}}{\partial p} = q(q-1)\big(\beta_{Eng} - \gamma_{Eng}qr\theta\big), \frac{\partial \dot{q}}{\partial q} = \Upsilon_1(2q-1) - q(q-1)\left(\frac{\alpha_{Eng}\theta}{q^2} + \gamma_{Eng}r\theta(p-1)\right)$$

$$\Upsilon_1 = (p-1)\big(\beta_{Eng} - \gamma_{Eng}qr\theta\big) + \frac{\alpha_{Eng}\theta}{q}, \Upsilon_2 = r\theta\gamma_{Op}^2 - \alpha_{Op} + \beta_{Op}(p-1)$$

Within the bounded strategy space $[0,1] \times [0,1]$ four boundary equilibria arise: $e_1 = (0,0)$, $e_2 = (0,1)$, $e_3 = (1,0)$ and $e_4 = (1,1)$. Substituting these points into the Jacobian under the realistic cost–benefit constraints listed in produces the results compiled in Table 2. The point e_1 is always a saddle because of the combined term. The node e_3 characterized by perfect compliance and no oversight, is likewise unstable since both diagonal elements of J are positive when compliance and supervision costs are non-zero. Points e_2 and e_4 may switch between stability and instability depending on the magnitudes of the short-term deviation benefit β_{Op} and the engineer rework penalty β_{Eng}; high values of these parameters make e_2 locally stable, whereas e_4 can become stable only when the deterrent γ_{Op} is large enough to offset the supervision cost α_{Eng}.

Table 2. Stability Analysis of the Boundary Equilibrium Points

Nash Equilibrium Points	J	$\det(J)$	$\mathrm{tr}(J)$	Stability
e_1	$\begin{bmatrix} -\alpha_{Op}-\beta_{Op} & 0 \\ 0 & \beta_{Eng}+\alpha_{Eng}\theta \end{bmatrix}$	$(-\alpha_{Op}-\beta_{Op})(\beta_{Eng}+\alpha_{Eng}\theta)$	$-\alpha_{Op}-\beta_{Op}+\beta_{Eng}$ $+\alpha_{Eng}\theta$	Saddle
e_2	$\begin{bmatrix} r\theta\gamma_1^2-\alpha_{Op}-\beta_{Op} & 0 \\ 0 & \alpha_2\theta-\beta_{Eng}+\gamma_{Eng}r\theta \end{bmatrix}$	$(r\theta\gamma_1^2-\alpha_{Op}-\beta_{Op})$ $(\alpha_{Eng}\theta-\beta_{Eng}+\gamma_{Eng}r\theta)$	$r\theta\gamma_1^2-\alpha_{Op}-\beta_{Op}$ $+\alpha_{Eng}\theta-\beta_{Eng}$ $+\gamma_{Eng}r\theta$	Stable or Unstable
e_3	$\begin{bmatrix} \alpha_{Op} & 0 \\ 0 & \alpha_{Eng}\theta \end{bmatrix}$	$(\alpha_{Op})(\alpha_{Eng}\theta)$	$\alpha_{Eng}\theta+\alpha_{Op}$	Unstable
e_4	$\begin{bmatrix} -r\theta\gamma_{Op}^2+\alpha_{Op} & 0 \\ 0 & \alpha_{Eng}\theta \end{bmatrix}$	$(-r\theta\gamma_1^2+\alpha_{Op})(\alpha_{Eng}\theta)$	$-r\theta\gamma_1^2+\alpha_{Op}$ $+\alpha_{Eng}\theta$	Stable or Unstable

An interior equilibrium $e_5=(p^*,q^*)$ exists if the cubic equation admits a root $q^*\in(0,1)$ and the corresponding $p^*\in(0,1)$ remains in the open unit interval. Under the inequalities assembled in the cubic has precisely one positive root, guaranteeing the presence of a single interior fixed point whenever $2\gamma_{Eng}rq^*\theta>\beta_{Eng}$. Evaluation of J at e_5 shows that the determinant becomes negative when $2\gamma_{Eng}rq^*\theta<\beta_{Eng}$, giving a saddle; if the opposite inequality holds, the determinant is positive, but the trace remains positive as well, yielding an unstable node. Hence the interior point is never evolutionarily stable. Figure 3 illustrates two representative phase portraits: in the first, no interior equilibrium exists, and the trajectories converge to the only stable boundary point e_2; in the second, e_5 appears but is unstable, so the system still lacks an ESS. The analysis therefore shows that passive evolutionary forces cannot maintain the ideal supervision–compliance synergy $(p=1,q=0)$ sustained managerial intervention is required to keep the system in a high-compliance, low-supervision regime.

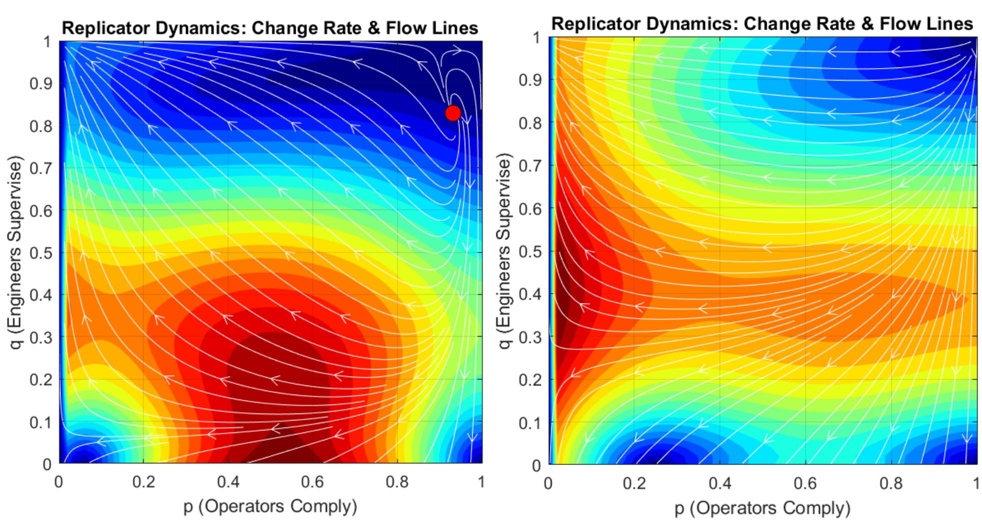

Figure 3. Phase diagram of two-player ECC game model

6. Conclusions

This study develops an evolutionary cooperation–competition game that explains how supervision and compliance co-evolve in human-centric manufacturing, and it couples that analytical model to a multi-sensor, AI-enabled perception stack capable of inferring real-time behavioral states. By merging evolutionary game theory with machine-learning–based compliance estimation, the work addresses a key Industry 5.0 question of how operator and engineer strategies adapt when decisions are driven by high-frequency data rather than static rules. The proposed ensemble Transformer fuses vision, tool signal, controller and interface streams through a Transformer architecture and produces continuous estimates of the shop-floor compliance rate. These estimates populate the payoff functions of the game and allow the replicator dynamics to evolve on live evidence, capturing the feedback loops that arise when supervision intensity responds to emerging deviations and, in turn, shapes operator behavior.

By integrating high-resolution sensing, deep-learning fusion and evolutionary modelling, the framework offers a practical decision-support tool for Industry 5.0 factories. It enables managers to diagnose behavioral risk in real time, simulate the impact of policy changes before implementation and deploy adaptive control strategies that preserve both human agency and production stability. This approach therefore advances the state of the art in sociotechnical systems engineering and provides a blueprint for embedding AI-powered behavioral analytics into future smart-manufacturing platforms.

References

Al-Amin, M., Tao, W., Doell, D., Lingard, R., Yin, Z., Leu, M. C., & Qin, R., Action Recognition in Manufacturing Assembly using Multimodal Sensor Fusion. Procedia Manufacturing, 39, 158-167, 2019.

Agote-Garrido, A., Martín-Gómez, A. M., & Lama-Ruiz, J. R., Manufacturing System Design in Industry 5.0: Incorporating Sociotechnical Systems and Social Metabolism for Human-Centered, Sustainable, and Resilient Production. Systems, 11(11), 2023.

Gudlin, M., Hegedić, M., Golec, M., & Kolar, D., Improving Time Study Methods Using Deep Learning-Based Action Segmentation Models. Applied Sciences, 14(3), 2024.

Bhattacharya, M., Penica, M., O'Connell, E., Southern, M., & Hayes, M., Human-in-Loop: A Review of Smart Manufacturing Deployments. Systems, 11(1), 2023.

McCoy, D. E., Game Theory as a Foundation of Evolutionary Psychology. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), Encyclopedia of Evolutionary Psychological Science (pp. 3309-3325). Springer International Publishing, 2021.

Renna, P., A Review of Game Theory Models to Support Production Planning, Scheduling, Cloud Manufacturing and Sustainable Production Systems. Designs, 8(2), 2024.

Zhang, Y., Wang, J., & Liu, Y., Game theory based real-time multi-objective flexible job shop scheduling considering environmental impact. Journal of Cleaner Production, 167, 665-679, 2017.

Liu, S., Li, L., Zhang, L., & Shen, W., Game-Based Collaborative Scheduling With Fuzzy Uncertain Migration in Cloud Manufacturing. IEEE Transactions on Automation Science and Engineering, 21(4), 7190-7202, 2024.

Tushar, W., Yuen, C., Saha, T. K., Nizami, S., Alam, M. R., Smith, D. B., & Poor, H. V., A Survey of Cyber-Physical Systems From a Game-Theoretic Perspective. IEEE Access, 11, 9799-9834, 2023.

Li, L., & Whang, S., Game Theory Models in Operations Management and Information Systems. In K. Chatterjee & W. F. Samuelson (Eds.), Game Theory and Business Applications (pp. 95-131). Springer US, 2021.

Gladysz, B., Tran, T.-a., Romero, D., van Erp, T., Abonyi, J., & Ruppert, T., Current development on the Operator 4.0 and transition towards the Operator 5.0: A systematic literature review in light of Industry 5.0. Journal of Manufacturing Systems, 70, 160-185, 2023.

Tian, R., Sun, L., Bajcsy, A., Tomizuka, M., & Dragan, A. D., Safety Assurances for Human-Robot Interaction via Confidence-aware Game-theoretic Human Models 2022 International Conference on Robotics and Automation (ICRA), 2022.

Baksi, R. P., Pay or Not Pay? A Game-Theoretical Analysis of Ransomware Interactions Considering a Defender's Deception Architecture. 2022 52nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks - Supplemental Volume (DSN-S), 2022.

Trist, E. L., & Bamforth, K. W., Some Social and Psychological Consequences of the Longwall Method of Coal-Getting. Human Relations, 4(1), 3-38, 1951.

Cherns, A., The Principles of Sociotechnical Design. Human Relations, 29(8), 783-792, 1976.

XIAO Renbin, Z. X., Research Progress of Opinion Polarization in Social Collective Behavior: Centered on Biased Assimilation and the Hostile Media Effects. Complex Systems and Complexity Science, 20(4), 1-9, 2023.

LIU Qi, X. R. An Opinion Dynamics Approach to Public Opinion Reversion with the Guidance of Opinion Leaders. Complex Systems and Complexity Science, 16(1), 13, 2024.

Assuad, C. S. A., Tvenge, N., & Martinsen, K., System dynamics modelling and learning factories for manufacturing systems education. Procedia CIRP, 88, 15-18, 2020.

Taylor, C., & Nowak, M. A., Transforming the dilemma. Evolution, 61(10), 2281-2292, 2007.

Song, M., Gong, X., Du, J., Lu, T., & Jiao, R. J., Population dynamics modeling of crowdsourcing as an evolutionary Cooperation-Competition game for fulfillment capacity balancing and optimization of smart manufacturing services. Computers & Industrial Engineering, 197, 2024.

Arboh, F., Dai, B., Quansah, P. E., Atingabilli, S., Drokow, E. K., & Addai-Dansoh, S., Safety first, but how? Examining the impact of safety leadership in frontline healthcare workers' safety performance during health crisis. Journal of Contingencies and Crisis Management, 32(3), 2024.

Yan, D., & Zhao, X., Improving Safety Compliance of Construction Workers: The Role of Safety Communication, Management Commitment to Safety, and Perceived Ease of Use. In J. Li, W. Lu, Y. Peng, H. Yuan, & D. Wang, Proceedings of the 27th International Symposium on Advancement of Construction Management and Real Estate Singapore, 2023.

Svertoka, E., Saafi, S., Rusu-Casandra, A., Burget, R., Marghescu, I., Hosek, J., & Ometov, A., Wearables for Industrial Work Safety: A Survey. Sensors (Basel), 21(11), 2021.

Indris, C., Ibrahim, F., Ibrahem, H., Bramesfeld, G., Huo, J., Ahmad, H. M., Hayat, S. K., & Wang, G., Supervised and Self-Supervised Learning for Assembly Line Action Recognition. J Imaging, 11(1), 2025.

Chen, C., Zhang, C., Wang, T., Li, D., Guo, Y., Zhao, Z., & Hong, J., Monitoring of Assembly Process Using Deep Learning Technology. Sensors (Basel), 20(15), 2020.

Huang, J., Wu, Y., Han, Y., Yin, Y., Gao, G., & Chen, H., An evolutionary game-theoretic analysis of construction workers' unsafe behavior: Considering incentive and risk loss. Front Public Health, 10, 991994, 2022.

Ji, P., Ma, X., & Li, G., Developing green purchasing relationships for the manufacturing industry: An evolutionary game theory perspective. International Journal of Production Economics, 166, 155-162, 2015.

Biographies

Mulang Song is a Ph.D. candidate in the School of Mechanical Engineering at the Georgia Institute of Technology (USA). He is graduating in Mechanical Engineering, with research that spans both mechanical and industrial engineering domains. He received both his B.S. and M.S. degrees in Mechanical Engineering with a focus on Computer Science, also from Georgia Tech. He has led and contributed to engineering projects through direct collaboration with industry, applying research in real-world production environments. His research interests include advanced manufacturing systems, intelligent production systems, and industrial AI applications. More info about his research: https://scholar.google.com/citations?user=jLL4nycAAAAJ.

Dr. Jiao is the editor-in-chief of Journal of Engineering Design and an associate professor of mechanical and industrial engineering at Georgia Tech, USA. Prior to joining the School of Mechanical Engineering at Georgia Tech in December 2008, he was an Assistant Professor and then Associate Professor in the School of Mechanical and Aerospace Engineering at Nanyang Technological University, Singapore. Before his career in Singapore, he was a Visiting Scholar in the Department of Industrial Engineering and Engineering Management at Hong Kong University of Science and Technology from 1998 to 1999. From 1993 to 1994, he was a Lecturer of Industrial Engineering in the School of Management at Tianjin University, China, and from 1988 to 1990, he worked as an Associate Lecturer in the Department of Industrial Design at Tianjin University of Science and Technology, China. More info about his research: https://scholar.google.com/citations?user=9yikEHAAAAAJ&hl=en&oi=ao.