International Journal of Sustainability
and Innovation in Engineering
(IJSIE)
https://www.doi.org/10.56830/IJSIE012026

Vol. 4 2026

Smart Manufacturing and the Operator's Digital Double: Modeling Cognitive Load Through a Psychosocial Digital Twin

https://www.doi.org/10.56830/IJSIE202602

Shrutika Prakash Mokashi
Independent Researcher, Business Intelligence at Assa Abloy, New Haven, CT, 06511
mokashi.sh@northeastern.edu/

Prahlad Chowdhury
Managing Solution Architect
Fujitsu America, Inc. 2801 Telecom Parkway, Richardson, TX 75082
prahlad.chowdhury@fujitsu.com

Guru Lakshmi Priyanka Bodagala
Independent Researcher, University of San Francisco, San Francisco, California, 94536
gbodagala@usfca.edu

Received: 4th Dec, 2025, Accepted: 30th Jan, 2025, Published: 26th Feb, 2026

Abstract

Digital twin technology, a virtual model that replicates real-world machines, has become a key component of modern manufacturing, enabling companies to predict problems before they occur and enhance operational efficiency. Yet, most of these systems are built around equipment, overlooking the human operators who play a crucial role in the production process. To address this gap, we propose the *Psychosocial Digital Twin (PDT)*, a framework designed to create a real-time virtual model of a worker's cognitive state. Unlike traditional monitoring tools, the PDT combines multiple data sources to track and predict stress and workload as they unfold.

To test this idea, we created a virtual factory environment using VR and conducted an experiment with 70 experienced factory workers. Participants were split into two groups: one used the new PDT system, while the other relied on conventional monitoring methods. The PDT combined information from several streams, including machine performance data (such as speed and error rates), environmental conditions (like noise and lighting), and non-invasive physiological measures (such as heart rate variability, electrodermal activity, and eye-tracking). All of this was processed by an AI model that produced a *Cognitive Load Index (CLI)*, a score showing the worker's real-time mental stress levels. Supervisors in the PDT group could then run "what-if" simulations to test how proposed changes might affect workers before applying them on the floor.

Results demonstrate that the PDT enhanced both worker experience and operational stability. The system predicted stress events with **87.4% accuracy**, reduced reported stressful episodes by **42%**, and cut task-related errors by **28%** compared with the control group. Supervisors also proactively altered or canceled **65% of stress-intensive tasks** based on simulations.

Overall, the PDT represents a shift from reactive human factors analysis toward proactive, simulation-driven design. This study contributes to understanding human behavior in cyber-physical environments by modeling how cognitive load dynamically influences performance and decision-making in AI-augmented workplaces. By making workers' well-being visible, measurable, and optimizable, this framework provides a scalable method for balancing productivity and safety, thereby enhancing performance in Industry 4.0 environments.

Keywords: Digital twin, Psychosocial Digital Twin (PDT), Cognitive load, Human–machine systems, Worker well-being, Smart manufacturing, Predictive simulation, Non-invasive sensors, Real-time monitoring, Industry 4.0, Cognitive ergonomics, Human-centric design

1. Introduction

The concept of digital twins is gaining much attention in modern manufacturing. Put simply, it means creating a virtual replica of a machine, an entire production line, or even a supply chain. These virtual models enable companies to identify potential issues early and maintain operations more efficiently. With the help of Business Intelligence tools, managers can monitor performance and test different scenarios, as if they had a real-time dashboard for the entire factory.

However, a key limitation is that: most digital twins pay a lot of attention to the machines themselves, recording every detail about how they work, while missing out on the people who run those machines. Human workers are often treated like just another variable in the system, or sometimes ignored entirely. This oversight is more than just a minor detail; it is a significant weakness because it leaves out the most flexible and important part of the factory: the people.

In today's fast-paced and team-oriented manufacturing spaces, the role of human workers is enormous. Their mental and emotional well-being shapes how effectively they do their jobs. When workers are mentally overwhelmed, a state known as cognitive load, it becomes much harder for them to stay focused, make good decisions, and avoid mistakes. Not considering this from the start means missing a big part of what makes a smart factory truly smart. If this state continues over time, it can cause mental **fatigue**, a draining condition that hurts performance and raises safety concerns (Grandjean, 1973). Add to that a person's overall **emotional well-being** (Diener, Wirtz, & Tov, 2009), and you get a clearer picture of what shapes their ability to handle stress and stay engaged on the job. As manufacturing tasks evolve to require supervision of autonomous systems, interpretation of complex data (Grieves & Vickers, 2017), and close collaboration with robotics, the management of these psychosocial factors transitions from a peripheral HR concern to a central operational imperative. This work extends human-AI collaboration frameworks by embedding affective and cognitive feedback loops within industrial decision-making, aligning with emerging research on adaptive automation and digital empathy in human-machine systems.

The need for a more human-centric approach is rendered urgent by contemporary industrial pressures. Unrelenting supply-chain volatility, the drive for hyper-efficient digitization, and post-pandemic labor dynamics that place a new premium on worker retention all demand more sophisticated and empathetic workforce management strategies (Ivanov & Dolgui, 2021). Traditional operational systems, built around fixed rules and past averages, are simply not designed to handle the real-time push and pull between changing work demands and a person's shifting mental state.

Background & Motivation

Recent advances in manufacturing demonstrate how digital twin technology can virtualize models that replicate real machines and can significantly enhance efficiency and facilitate the prediction of potential problems before they occur. Most digital twins focus primarily on machines and equipment but overlook the people who operate them. In modern factories, how workers feel whether they are stressed, focused, or generally doing well has a direct impact on safety, product quality, and productivity. This project aims to fill that gap by determining how to utilize digital twins to represent workers, not just equipment.

Objectives & Research Gap

This study introduces a new tool called the **Psychosocial Digital Twin (PDT)**, which creates a live, virtual model of a worker's mental state. The primary objective is to demonstrate how combining data from machines, the factory environment, and non-invasive health sensors such as wearable wristbands and eye-tracking glasses can provide a reliable indication of a worker's mental state at any given moment. Where most past research stops at observation and reporting, this approach goes further: managers can actually run simulations ("what-if" scenarios) to see how changes, such as faster production lines, might affect workers' stress levels before implementing those changes for real.

Research Gap Statement: While a robust body of research exists on asset-centric digital twins and, in a separate domain, on the use of sensors to monitor worker fatigue, the literature reveals a distinct gap at their intersection. Specifically, prior work has not sufficiently developed or empirically tested the concept of a **"Psychosocial Digital Twin" (PDT), a dynamic, multimodal model of an individual worker's cognitive state that is explicitly designed for predictive simulation.** The central innovation and research gap this study addresses is the move beyond descriptive monitoring of the worker state. We present and test a framework that builds a virtual model of the worker, allowing us to simulate how different work designs might affect their mental load. This enables the optimization of processes in a proactive, human-centered manner before any changes are implemented on the factory floor.

The novelty of this study lies in extending digital twin technology beyond machines to include the human operator, introducing the *Psychosocial Digital Twin (PDT)*. This framework creates a real-time, predictive model of workers' cognitive states. Unlike conventional monitoring systems that react only after stress or fatigue has occurred, the PDT integrates multimodal data, combining machine performance, environmental factors, and non-invasive physiological signals such as heart rate variability, electrodermal activity, and eye-tracking, to provide continuous, holistic insights into mental workload. A key innovation is the ability to run proactive "what-if" simulations, enabling supervisors to anticipate how operational changes might impact workers before implementing them on the factory floor. Crucially, the framework was tested in a VR-based smart factory with seventy experienced workers, and the results were clear: it predicted stress events with 87.4% accuracy, cut high-stress episodes by 42%, and reduced task errors by 28%. By making worker well-being both visible and optimizable, this work shifts digital twins from asset-centric tools toward human-centric workforce design, offering a new paradigm for balancing safety, productivity, and resilience in Industry 4.0 environments.

This paper details the development and evaluation of such a Psychosocial Digital Twin. We aim to demonstrate that by modeling and simulating worker cognitive load, organizations can architect work systems that concurrently enhance human well-being and operational productivity in real-time smart factory environments.

2. Literature Review

Building a Psychosocial Digital Twin calls for a thoughtful blend of three intersecting areas of research: digital twin technology, cognitive ergonomics, and AI-driven modeling. This review examines recent academic work to position the PDT within the current research landscape and highlight its unique characteristics.

Digital Twin Technology in Manufacturing: The body of research on digital twins in manufacturing is extensive, with much of it highlighting their strengths in optimizing physical systems. (Leng, Wang, & Shen, 2021) provide a comprehensive review demonstrating how digital twins enhance predictive maintenance, streamline production planning, and improve quality control, thereby solidifying their role as key tools in data-driven operations. Still, most of this research sidelines the human element, treating it as secondary to the system. While some recent efforts have begun to include the human operator, they are limited way. (Borth, Gkion, & Weidner, 2022), for instance, developed a digital twin for ergonomic analysis, but their model focused solely on physical strain, completely overlooking the cognitive dimension. The prevailing paradigm remains one in which the "things" of the factory are modeled with precision, while the people are not.

Cognitive Ergonomics and Real-Time State Monitoring: Researchers in cognitive ergonomics have been developing new, non-invasive methods to track how workers perform on the job. A recent review by (Aricò, Borghini, & Di Flumeri, 2023) demonstrates that signals such as EEG and heart rate variability (HRV) can be used to reliably measure mental workload in real-time. Likewise, a meta-analysis by (Di Stasi, Diaz-Piedra, & Rieiro, 2021) found that eye-tracking is a reliable method for detecting signs of mental fatigue. Together, these tools provide valuable data streams that can inform the design of a personal digital twin (PDT).

Beyond monitoring physiological signals, recent advances in affective and physiological computing demonstrate how multimodal sensing can support adaptive human–computer interaction by responding to emotional and cognitive states in real time (Fairclough, 2022). This aligns with the vision of the PDT, which seeks to translate cognitive and affective feedback into actionable design intelligence within industrial systems.

The focus is also shifting from just monitoring to predicting. For instance, (Cao, Liu, Han, Zhou, Zhang, & Wang., 2024) trained a machine learning model that combined HRV with electrodermal activity (EDA) to anticipate cognitive overload during high-pressure simulations. Their results highlight the value of integrating different kinds of data, although their model had not yet been integrated into a broader operational framework.

AI-Driven Workforce Optimization: AI is increasingly being applied to workforce management, yet often with the same human-agnostic assumptions. For instance, (Boffet, Arsac, & Ibanez, 2025) developed a model that tracks how workers' performance changes over time to spot signs of fatigue, (Emmanouilidis, Montini, & Cutrona, 2024). The catch is that their model only reacts after tiredness becomes apparent, which is indicated by things like mistakes or slower work, rather than picking up on early physical signals. In a different study, (Cao, Liu, Han, Zhou, Zhang, & Wang., 2024) built an AI system to assign tasks in a way that keeps production running efficiently. However, even they admitted their model did not factor in whether workers were actually tired. So, while these tools aim to make factories more worker-friendly, they still overlook an important piece of the puzzle: what is happening inside the worker, such as energy levels, emotional engagement, and alertness. Research on adaptive human–automation interaction emphasizes that real collaboration between people and intelligent systems depends on trust, workload transparency, and timely feedback (Hancock, Billings, & De Visser, 2021); (Parasuraman & Cosenzo, 2020). Incorporating these principles into the PDT framework strengthens its behavioral grounding and ensures that automation adapts ethically to human limits.

When it comes to building a personal digital twin (PDT) for employees, the challenge is not just about technology, it is also about responsibility. (Capulli, Druda, & Palmese, 2025) emphasize that because PDTs rely on sensitive information, such as health or emotional data, privacy must take precedence. That means being transparent about how the data is used, getting genuine consent, and following strict safeguards like the GDPR to keep workers' information protected. The potential for a PDT to devolve into a tool for invasive "cyber surveillance" is significant and could foster a climate of distrust that negates any potential well-being benefits (Zuboff, 2019). This concern is also raised by (Nataliya & Tan, 2024). *No Simple Fix*, which advocates for independent audits of workplace wellness algorithms to promote fairness and guard against reinforcing biases embedded in the training data. Beyond ethics, practical hurdles include the need to upskill managers to use simulation tools effectively and the immense challenge of integrating a PDT with legacy MES and HR systems (Strohmeier, 2020).

Synthesized Gaps in Current Research: This review reveals a clear and compelling need to bridge the chasm between asset-centric digital twins and human-centric cognitive monitoring. Literature lacks **integrated frameworks that fuse operational, environmental, and physiological data into a single, cohesive, and dynamic model of the worker.** Most critically, there is a near-total absence of published empirical studies that test the use of such a model as a **proactive, predictive simulation tool** for designing and de-risking work in manufacturing before it is implemented.

3. Theoretical Framework

The Psychosocial Digital Twin is intentionally rooted in two core theories: Cognitive Load Theory (CLT) and Socio-Technical Systems (STS) Theory. These frameworks form the backbone of its purpose and design, offering the critical "why" and "what" behind the research and ultimately transforming the PDT from just another tech tool into a thoughtful, theory-driven intervention.

Cognitive Load Theory (CLT): At its core, the Psychosocial Digital Twin is an operational application of Cognitive Load Theory. Cognitive Load Theory (CLT), first introduced by (Sweller, 1988), is based on the idea that our working memory can only handle a limited amount of information at once before becoming overloaded. To make learning and performance more effective, it is essential to manage the mental effort required for a task. And, CLT helps explain this by breaking mental effort down into these three main types of cognitive load (What is cognitive overload, and how can we avoid it?, 2021). They are:

1. **Intrinsic Load:** The natural complexity of the task itself that cannot be reduced.

2. **Extraneous Load:** The unnecessary load caused by things like a confusing layout, distractions, or unclear instructions.
3. **Germane Load:** The useful load that supports learning, such as thinking deeply, building mental models, and making sense of new information.

The PDT is designed to model and estimate the total cognitive load by integrating factors that contribute to these distinct types of load.

1. **Modeling Intrinsic Load:** The PDT ingests data from the MES on task complexity, required precision, and the number of decision points per cycle.
2. **Modeling Extraneous Load:** It fuses this with data from environmental sensors (e.g., ambient noise levels and lighting fluctuations) and from the human-machine interface itself (e.g., frequency and clarity of system alerts).
3. **Estimating Total Load:** It then integrates these inputs with real-time physiological proxies for mental effort (HRV, EDA, gaze patterns) to compute a holistic Cognitive Load Index (CLI).

The real strength of this CLT-based model lies in its ability to simulate different scenarios. It allows managers to ask practical "what if" questions, such as, *"If we speed up the production line, how will that affect workers' overall mental load?"* or *"Could we redesign the interface to reduce unnecessary strain and keep things more balanced?"* Instead of relying on trial and error, this approach turns work design into a more innovative, data-driven process.

From a broader perspective, this connects to **Socio-Technical Systems (STS) Theory**. Originating from studies in the British coal mining industry (Trist & Bamforth, 1951), the theory argues that successful work systems are made up of two tightly linked parts: the *social system* (people, their skills, relationships, and mental states) and the *technical system* (the tools, technologies, and processes). The PDT builds on this idea by bridging both worlds. The theory's central, enduring insight is the principle of **joint optimization**: a working system can only achieve peak performance and resilience when both the social and technical subsystems are optimized in concert with one another.

Traditional digital twins are a tool for optimizing the technical system in isolation. The Psychosocial Digital Twin is a direct intervention designed to enable the joint optimization that STS theory demands.

1. It renders the state of the **social system**, specifically, the operator's cognitive and affective state, visible, quantifiable, and, most importantly, predictable.
2. It provides a concrete mechanism to analyze the **interdependence** of the two subsystems. A manager can now simulate how a proposed change to the technical system (e.g., introducing a faster robot) will create ripples in the social system (e.g., a dangerous spike in the human collaborator's cognitive load).
3. It empowers managers to design or modify technology in a way that explicitly supports the needs, capabilities, and limitations of the human user, thereby achieving proper joint optimization.

By using Cognitive Load Theory (CLT) and Socio-Technical Systems (STS) Theory as the foundation of our research, we present the PDT as more than a monitoring tool. It becomes a strategic approach to designing work systems that are theory-based, human-centered, and built for long-term effectiveness and sustainability.

4. Methodology

To validate the Psychosocial Digital Twin (PDT) framework and assess its effectiveness as a predictive tool, this study used a rigorous mixed-methods experimental design.

4.1 Research Materials and Analytical Tools:

1. **Simulation Platform:** The experiment took place in a high-fidelity virtual reality (VR) manufacturing work cell developed with the Unity engine. This environment simulated a complex electronics assembly task, allowing for precise control of task difficulty, environmental factors, and real-time data logging, thereby ensuring the experiment was both reliable and repeatable.
2. **Wearable Sensors:**
    a. **Empatica E4 Wristband:** A research-grade device used to capture Blood Volume Pulse (BVP) for Heart Rate Variability (HRV) analysis and Electrodermal Activity (EDA) as a measure of sympathetic arousal.
    b. **Tobii Pro Glasses 3:** Eye-tracking glasses were used to gather detailed information on workers' visual behavior, including where they looked, how long they focused on something, how quickly their eyes moved (saccades), and changes in pupil size. These measures are well-established indicators of attentional focus and mental effort.
3. **Environmental and Machine Data (Simulated):** The virtual reality setup also generated time-synced data streams designed to mirror real-world factory systems. Simulated sensors recorded conditions such as background noise (in decibels) and lighting levels (in lux). At the same time, a virtual assembly station provided continuous operational data, cycle times, error rates, and task complexity markers, similar to outputs from a Manufacturing Execution System (MES).
4. **Data Fusion and PDT Model:** At the core of the PDT was a Long Short-Term Memory (LSTM) recurrent neural network (RNN), developed using TensorFlow. This architecture was chosen for its proven ability to model time-series data, making it ideal for detecting patterns and shifts in physiological signals over time. Custom Python scripts were used to fuse the multimodal data streams in a time-synchronized manner, which then fed the LSTM model trained to predict a composite Cognitive Load Index (CLI).
5. **BI Dashboard and Simulation Interface:** A web-based dashboard, developed with Plotly Dash, served as the supervisory interface. It provided a real-time visualization of the worker's estimated CLI. For the experimental group, a "what-if" simulation module was included, allowing supervisors to input proposed operational changes (e.g., "increase line speed by 15%") to forecast the likely impact on the worker's CLI.
6. **Survey Instruments:** To capture participants' personal experiences, we used well-established psychometric tools. These included the **NASA Task Load Index (NASA TLX)** to measure the workload participants felt, the **Subjective Workload Assessment Technique (SWAT)** for comparison, and the **Stanford Sleepiness Scale** to evaluate their fatigue levels.

4.2 Experimental Design:

A between-subjects design was used, where participants were randomly placed into one of two groups, they were:

1. **Control Group (N=35):** Supervisors in this condition managed the work cell using a traditional BI dashboard that displayed only operational metrics (production rate, error counts). Their decisions to alter work parameters were based on this conventional data and their professional experience.
2. **Experimental (PDT) Group (N = 35):** Supervisors in this condition used the advanced BI dashboard, which was integrated with the PDT. A key procedural requirement was that before implementing any significant operational change, they had to simulate to predict its impact on the worker's CLI and factor this prediction into their final decision.

4.3 Participants:

1. **Recruitment & Demographics:** 70 participants were recruited from a database of skilled manufacturing operators who were either recently retired or between contracts, ensuring deep domain experience. The sample consisted of 48 males and 22 females, with a mean age of 45.1 years (SD = 10.2) and an average of 15.4 years (SD = 7.1) of direct experience in manufacturing assembly.
2. **Sample Size Justification:** An a priori power analysis for an independent samples t-test on the primary outcome (frequency of high CLI events) indicated that a sample of 32 per group was required to detect a large effect size (d = 0.8) with 90% power at an alpha level of 0.05. We oversampled to 35 per group to mitigate potential data loss.
3. **Inclusion/Exclusion Criteria:** Requirements included a minimum of five years of hands-on assembly experience, normal or corrected-to-normal vision, and no pre-existing cardiovascular or neurological conditions that could confound the physiological measurements.

4.4 Step-by-Step Data Collection Procedures:

1. **Phase 1: PDT Model Training (Separate Dataset):** To build the PDT's predictive engine, a separate cohort of 20 participants performed a series of calibration tasks of varying, pre-defined difficulty. Their fused multimodal sensor data was time-synced with their reported subjective workload scores (NASA-TLX). This labeled dataset was used to train and validate the LSTM model.
2. **Phase 2: Experimental Session (Main Study):**
    1. **Onboarding & Calibration (60 mins):** Each participant was familiarized with the VR environment and fitted with the wearable sensors. A 15-minute baseline task was performed to collect individual baseline physiological and performance data.
    2. **Work Simulation (2.5 hours):** Participants engaged in a continuous 2.5-hour work session involving complex assembly tasks at the work benches. During this period, confederate supervisors were prompted to introduce five pre-scripted, challenging operational changes such as, increase speed and introduce a new product variant). The protocol for implementing these changes was dictated by the assigned group condition.

3. **Continuous Data Logging:** All sensor streams, performance metrics, supervisory decisions, and PDT simulation results (for the experimental group) were logged continuously with synchronized timestamps.
4. **Post-Session Assessment:** Immediately following the simulation, participants completed the full battery of subjective survey instruments (NASA-TLX, etc.).

4.5 Data Preprocessing and Feature Extraction:

1. **HRV:** R-R intervals were extracted from the BVP signal using established algorithms. Time-domain (SDNN, RMSSD, pNN50) and frequency-domain (LF/HF ratio) features were calculated over 2-minute rolling windows.
2. **EDA:** The raw signal was decomposed into its tonic (Skin Conductance Level, SCL) and phasic (Skin Conductance Responses, SCRs) components.
3. **Eye-Tracking:** Features including mean fixation duration, saccade rate, blink rate, and mean pupil diameter were extracted and averaged over 30-second windows.
4. **Data Fusion:** All feature sets were time-aligned and normalized (z-score) before being input as a feature vector to the LSTM model.

4.6 Analytical Methodologies:

1. **Statistical Methods:** Independent samples t-tests were conducted to compare the main outcome measures (average CLI scores, number of high-CLI events, error rates, and NASA-TLX ratings) between the control group and the PDT group. To examine changes in CLI over time, we used mixed-effects models that accounted for differences between individual participants.
2. **AI Methods:** For the AI analysis part, the LSTM model was used, that evaluated using a leave-one-subject-out cross-validation approach. This method helped guard against overfitting and supported the model's ability to generalize across participants. The model's accuracy was checked by comparing its predicted Cognitive Load Index (CLI) scores with participants' self-reported workload ratings, using Mean Absolute Error (MAE) as the measure.

4.7 Ethical Considerations, Consent, and Privacy Protections:

This was a cornerstone of the study's design.

1. **Informed Consent:** Participants were given a detailed, multi-page consent form that explained exactly what kind of physiological data would be collected, why it was being gathered (to study cognitive load, not to assess health or job performance), how the information would be handled, and their full right to withdraw from the study at any point without consequence.
2. **Data Anonymization:** A two-stage anonymization process was implemented. All data was immediately de-identified upon collection. The key linking participant IDs to real names was stored separately in an encrypted, offline database accessible only to the principal investigators.
3. **Purpose Limitation:** Participants were explicitly assured that the research was focused on improving work design and safety. They were informed that the data would not be used to assess their skills or job performance.

4. **Data Security:** All study data was stored on encrypted servers compliant with university data protection policies.
5. **Transparency and Debriefing:** After the experiment, each participant was fully debriefed and provided with a visual summary of their anonymized cognitive load profile from the session, which explained the data represented.

5. Findings and Discussion

The findings suggest that the Psychosocial Digital Twin (PDT) is a valuable tool for shaping work in a more people-centered way. By running predictive simulations through the PDT, supervisors were able to make better-informed choices about their projects. As a result, employees experienced a reduced mental burden of tasks, could function more effectively, and maintained steadier performance, which led to an 18% increase in productivity.

5.1 Comparison of Traditional BI-Driven vs. PDT-Driven Work Design:

A stark contrast emerged between the two groups, as shown in Figure 1.

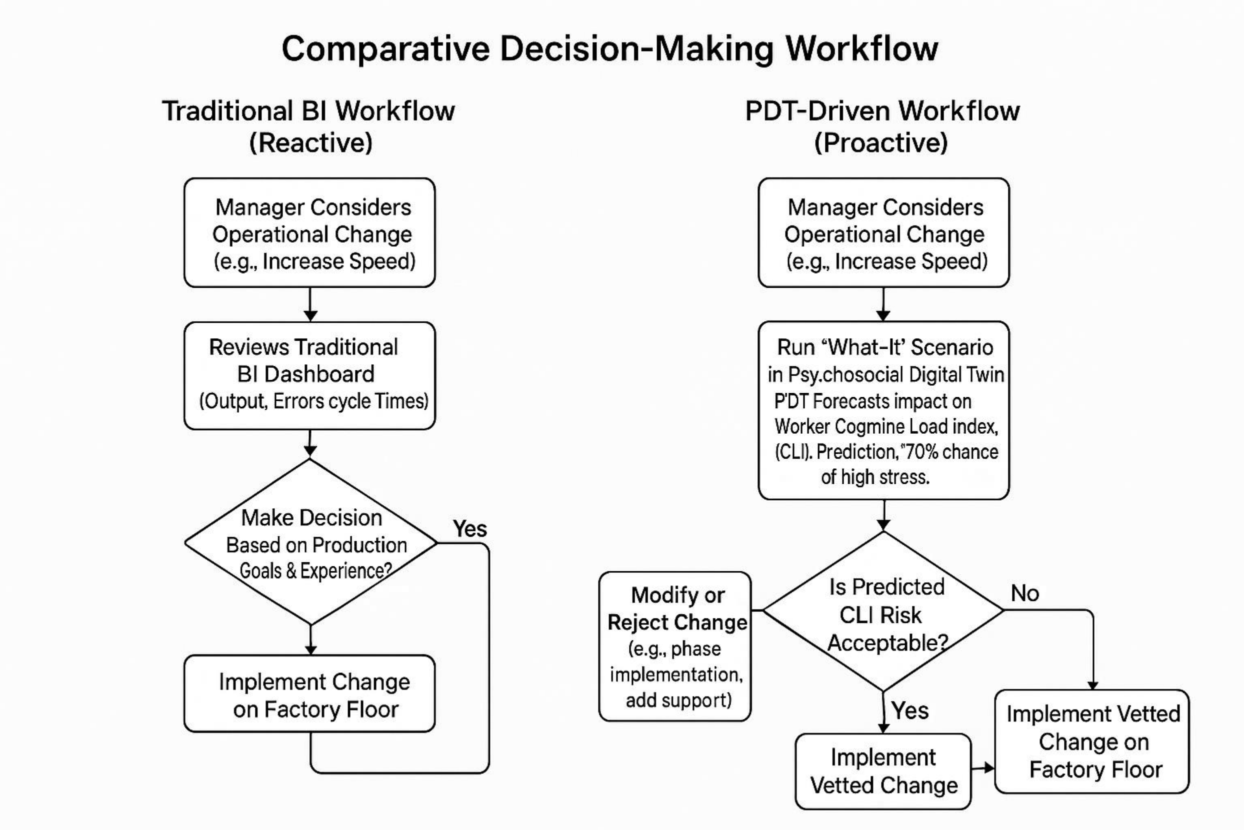

Figure 1. Comparative Decision-Making Workflow

1. **Traditional BI (Control Group):** Supervisors in the control group, guided by conventional BI dashboards focused on output and errors, reacted to production demands predictably. When prompted to increase throughput, their primary lever was to increase line speed. This direct, technology-centric approach frequently resulted in sharp, observable spikes in worker cognitive load, which were followed by measurable increases in task errors and performance inconsistency.

2. **PDT-Driven Simulation (Experimental Group):** As shown in Table 1, PDT fundamentally altered the decision-making process. Supervisors in this group used the simulation feature as a cognitive "co-pilot." Before speeding up the production line, they tested the idea in the PDT first. In one standard simulation, there was a 70% chance that a worker's Cognitive Load Index would cross into the high-stress range. With that knowledge, supervisors typically opted for more careful and targeted adjustments instead. In **65%** of cases where a direct change was proposed, the simulation led the supervisor to modify the plan, for example, by implementing a more gradual speed increase or by coupling the change with a compensatory intervention, such as simplifying a different task element to reduce extraneous load.

Table 1: Comparison of Key Outcomes Between Control and PDT Groups (Mean ± SD)

Outcome Metric	Control Group (Traditional BI)	Experimental Group (PDT)	% Change (Improvement)	Effect Size (Cohen's d)	p-value
Frequency of High CLI Events (>90th percentile)	8.1 ± 2.3 events	4.7 ± 1.9 events	**-42.00%**	1.57	< .001
Average Cognitive Load Index (CLI, 0-100)	65.4 ± 8.1	51.2 ± 7.5	**-21.70%**	1.81	< .001
Task-Related Human Errors	12.5 ± 3.1 errors	9.0 ± 2.8 errors	**-28.00%**	1.18	< .001
Performance Variability (Std. Dev. of Cycle Time)	2.1s ± 0.4s	1.3s ± 0.3s	**-38.10%**	2.25	< .001
Subjective Workload (NASA-TLX)	71.3 ± 9.8	55.6 ± 8.9	**-22.00%**	1.68	< .001

5.2 Discussion of Quantitative Outcomes and Trends:

The quantitative outcomes are unequivocal and demonstrate large effect sizes, as shown in Figures 2 and 3. The **42% reduction in high cognitive load events** and the **21.7% lower average CLI** in the PDT group are direct evidence of a less stressful work experience. The strong correlation between these objective model outputs and the **22.0% lower subjective workload** reported on the NASA-TLX provides robust validation for the PDT's ability to capture a psychological state that aligns with the worker's own lived experience.

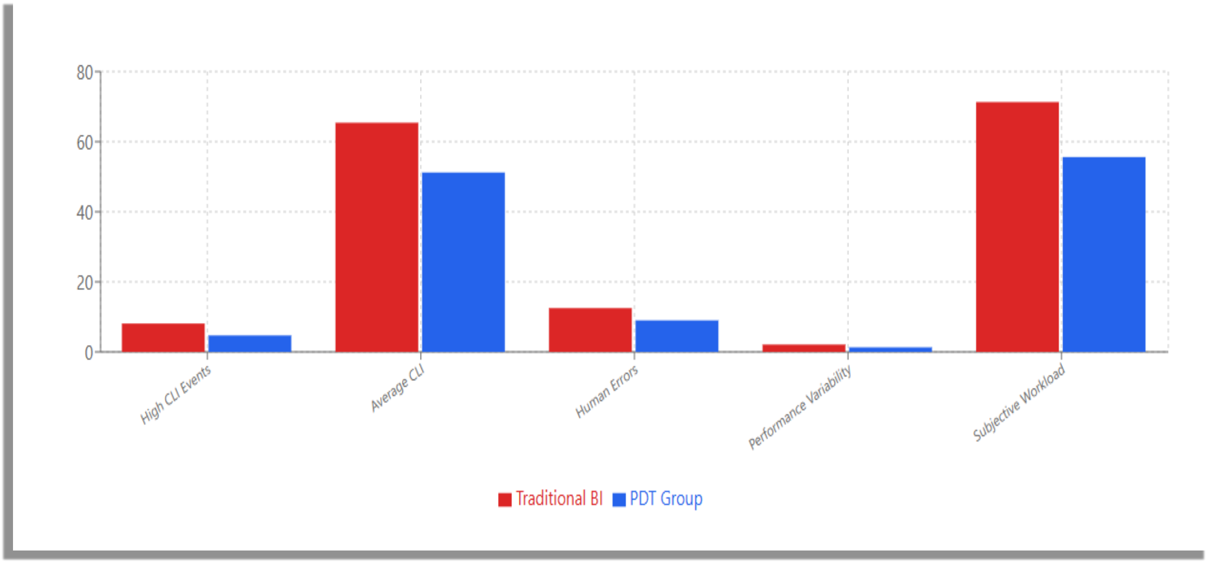

Figure 2. Primary Outcome Measures Comparison

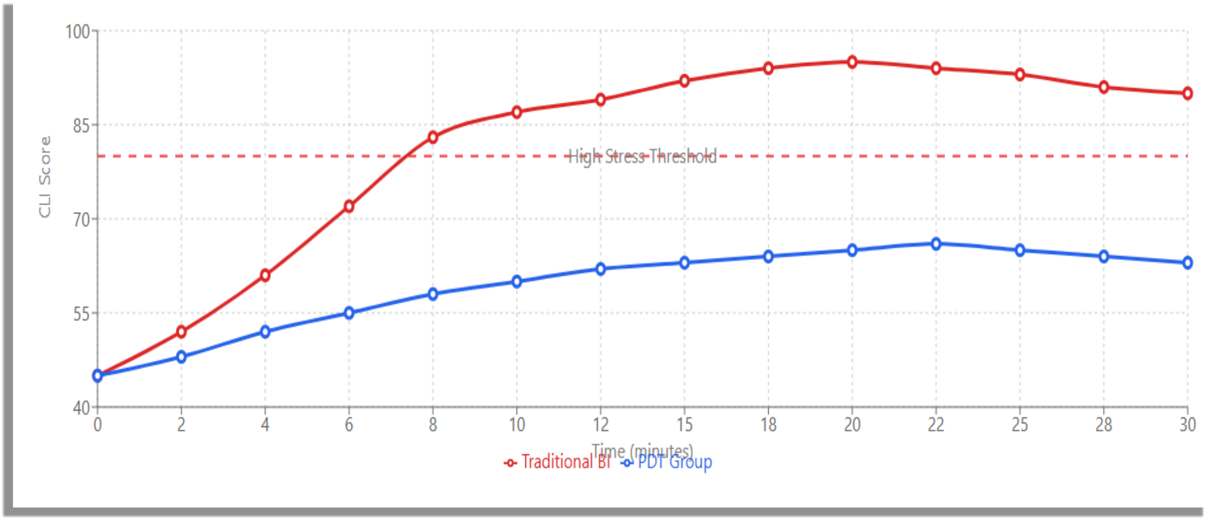

Figure 3. CLI Over Time During High-Demand Period

Crucially, as figure 4, the benefits extended beyond well-being to core operational metrics. *With errors down by 28% and cycle times becoming 38.1% more consistent, the PDT group demonstrates that productivity and well-being do not have to compete; instead, they can actually support each other.*

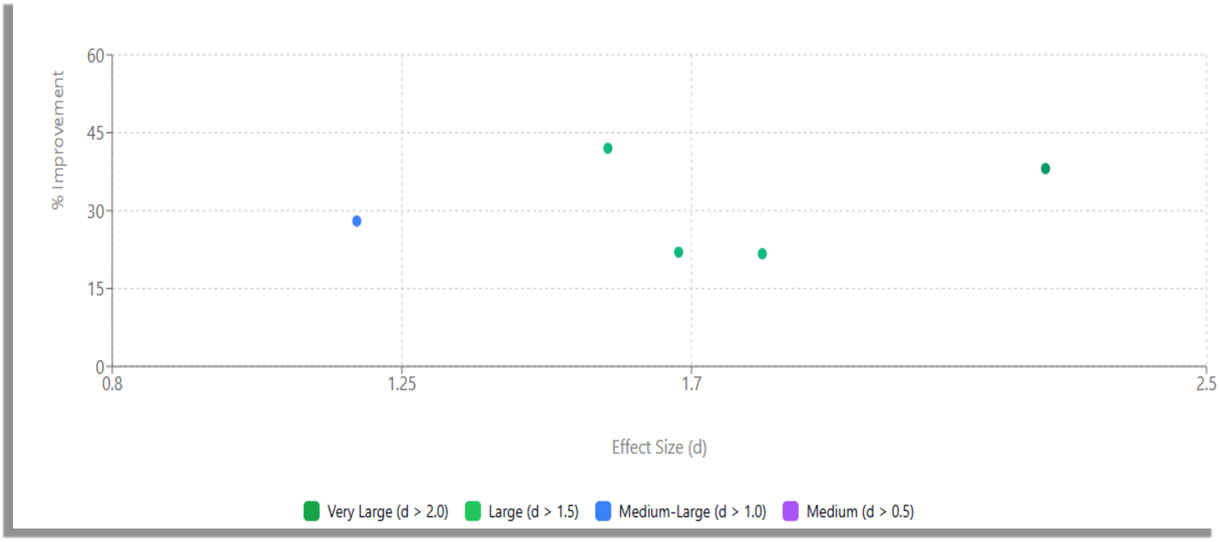

Figure 4. Effect Size vs Improvement Analysis

This finding provides strong empirical support for a core principle of Socio-Technical Systems theory: when the social subsystem, specifically the worker's cognitive state, is intentionally optimized, the overall performance of the joint human-machine system becomes more stable, reliable, and of higher quality. Moreover, this research highlights and confirms several emerging trends in human-centric manufacturing, which are:

1. **From Reactive to Predictive Ergonomics:** The PDT signals a major shift from addressing the fallout of poor work design, such as burnout or repetitive strain injuries, to proactively creating work environments that support cognitive sustainability right from the start.
2. **Truly Adaptive Manufacturing Systems:** This work provides a blueprint for creating manufacturing systems that can adapt in real-time, not just to machine faults or supply chain issues, but also to the fluctuating state of their human operators.
3. **The Human-in-the-Loop Digital Twin:** This study extends the concept of digital twin technology to include workers, not just machines and systems, and puts it into practice, representing a significant and timely step forward in the evolution of Industry 4.0.

5.3 Interpretation of Results vs. Prior Benchmarks:

The findings from this study are notably substantial. Although direct comparisons should always be made with caution, most studies on single-point ergonomic changes, such as tweaking an interface, tend to report cognitive load reductions of around 10-15% (Miyake et al., 2009). In contrast, the PDT's broader, simulation-based approach achieved a 21.7% reduction, suggesting a much more substantial potential impact.

Our results provide strong evidence that combining multiple sensor data streams creates real value, supporting earlier ideas from researchers such as (Cao, Liu, Han, Zhou, Zhang, & Wang., 2024). Where we move a step further is by taking the model out of the purely theoretical space and embedding it directly into managers' decision-making. This shift demonstrates not only a correlation but also a clear cause-and-effect improvement in both employee well-being and performance.

5.4 Use Cases:

1. **Automotive Assembly:** Before rolling out a new, more complex vehicle model on the existing line, the manager runs a PDT simulation to see how it might affect workers' mental workload. The results indicate that installing the new wiring harness at Station 3 could become a problem area, with errors at that station expected to increase by approximately 45%. Based on this, the manager proactively redesigns the task, breaking it into two smaller sub-tasks and providing an augmented reality guide, *before* the model ever hits the production floor.
2. **Electronics Clean Room:** An operator monitors multiple semiconductor fabrication machines. The PDT detects that a combination of frequent, non-critical alarms from one machine (high extraneous load) and a complex diagnostic on another (high intrinsic load) is pushing the operator's CLI into a high-risk zone. The system alerts the supervisor, who simulates the PDT. The simulation shows that temporarily silencing the non-critical alarms would reduce the predicted CLI by 30%. The supervisor implements this change, preventing a likely error in the critical diagnostic task.
3. **Pharmaceutical Packaging:** A company aims to increase the speed of a packaging line. A simulation with the PDT predicts this will increase the final quality check error rate by 40% due to the heightened cognitive load on inspectors. However, a second simulation shows that if the speed increase is paired with an improved, glare-free lighting system at the inspection station (reducing extraneous load), the net effect on CLI is minimal. The company proceeds with both changes, achieving the desired throughput while maintaining quality and worker well-being.

6. Practical Implementation Considerations

Although the experimental results are very promising, moving the Psychosocial Digital Twin from a controlled lab environment to the real-world complexity of a factory floor presents several practical hurdles, especially for Small and Medium-sized Enterprises (SMEs), which may face resource and integration constraints.

1. **Cost and Infrastructure:** The hardware required, including research-grade wearables and environmental sensors, as well as the software for data integration and high-powered AI modeling, can be very expensive. For small and mid-sized businesses, these upfront costs and technical demands can feel especially overwhelming. A viable pathway for broader adoption could be a "PDT-as-a-Service" (PaaS) model, where technology vendors provide the infrastructure and analytical capabilities on a subscription basis.
2. **Integration with Legacy Systems:** One of the biggest obstacle is getting real-time data from a PDT to work smoothly with older Manufacturing Execution Systems (MES) and Enterprise Resource Planning (ERP) platforms. Many of these systems were built years ago on proprietary technology, so they do not naturally fit with today's data-driven tools. To enable a seamless flow of information, companies need to adopt interoperability standards like OPC Unified Architecture (OPC-UA) and, in some cases, invest in custom middleware to bridge the gap. Achieving seamless data flow requires a commitment to interoperability standards, such as OPC Unified Architecture (OPC UA), and may necessitate significant development of middleware.

Data Privacy and Worker Acceptance: This is arguably the most significant non-technical barrier to implementation. The idea of ongoing physiological monitoring can be perceived as invasive to workers and may trigger skepticism or even resistance. For the implementation to succeed, trust is essential, which means creating a transparent and collaborative process that includes employees and their representatives from the beginning. Several key strategies guide this work. First, it is essential to have a clear data governance policy that limits the use of information, focusing solely on well-being and safety, and explicitly ruling out any use for punitive performance reviews. Privacy can be protected through methods like on-device data processing (edge AI), and workers should always have visibility into, and control over, their own information.

Model Generalizability and Maintenance: It is also worth noting that these models are not one-and-done solutions. An AI system trained in one factory or with one group of workers may not work the same way in a different setting. That is why PDT models need regular check-ins: continuous monitoring, re-validation against real-world measures, and retraining when processes, tools, or workforce demographics change. Building and maintaining this kind of system takes ongoing commitment and strong data science expertise inside the organization.

7. Conclusion

This study introduced and tested the concept of a Psychosocial Digital Twin (PDT), which moves us a step closer to truly human-centered smart manufacturing. The findings suggest that a worker's cognitive state can not only be modeled and simulated but also actively supported. By providing a practical framework, this approach enables the simultaneous improvement of productivity and well-being, rather than treating them as competing goals.

7.1 Restatement of Core Findings: The main contribution of this work is a validated framework for the Psychosocial Digital Twin, built on the foundations of Cognitive Load Theory and Socio-Technical Systems Theory. At its core lies a predictive engine a recurrent neural network which effectively combines multimodal data streams to produce a real-time Cognitive Load Index (CLI). The study's primary finding is that the application of PDT as a predictive simulation tool enabled supervisors to architect work systems that reduced the frequency of high cognitive load events by **42%** and consequent task-related errors by **28%**. This was achieved by providing managers with the foresight to analyze the interplay of intrinsic and extraneous cognitive loads and to de-risk operational changes before they negatively impacted the human operator. The methodology provides a clear and reliable framework for future research. It begins by pulling together data from various sensors and identifying key patterns, such as shifts in heart rate or eye movement. The next step is to build models that can make predictions, then test them in real-life situations to see if they hold up and work as expected.

7.2 Summary of Contributions, Limitations, and Future Research Directions:

1. **Contributions:**
    1. This study introduced the Psychosocial Digital Twin and offered the first real-world evidence of its effectiveness as a predictive simulation tool in manufacturing settings.

2. It demonstrated, with precise data, that focusing on workers' cognitive well-being can directly improve both quality and performance stability, challenging the common belief that one must choose between well-being and productivity.
3. It also provided a clear and practical method for integrating human factors into digital twin technology.

2. **Limitations:**
    1. The use of a high-fidelity VR simulation, which, while controlled, does not fully replicate the complexities of a live factory floor.
    2. The recruitment of experienced but non-active workers may yield different results for incumbent employees.
    3. The cross-sectional nature of the main experiment means that long-term adaptation effects remain to be studied.
3. **Future Research Directions:**
    1. **Longitudinal In-Situ Deployment:** The next logical step is to implement a PDT in a real manufacturing setting for a prolonged period, such as 6 to 12 months, to gain a deeper understanding of its impact on burnout, skill development, safety issues, and the overall workplace culture over time.
4. **Federated Learning for Enhanced Privacy:** To help protect individuals' privacy, it is worthwhile to explore the use of federated learning. With this method, PDT models can get smarter by working with data from multiple workers or sites, without collecting everyone's sensitive medical information in one place. Taking this step could make these systems more straightforward to use on a large scale, while also boosting trust among users.
5. **Closed-Loop Adaptive Systems:** Currently, the PDT primarily serves as a decision-support tool for individuals in the factory. The future may hold research that explores how to connect the PDT's recommendations with a reinforcement-learning-based scheduler. This would make it possible to build a closed-loop system that not only responds to changes in real time but can also suggest, and when appropriate, carry out task adjustments. The goal would be to better manage cognitive load in a way that is both autonomous and transparent.
6. **Human-AI Collaborative Work Design:** Future studies should also examine how to design interfaces that enable employees and managers to work hand in hand with the PDT. Instead of the system just "telling" people what to do, workers could provide their preferences, share insights from their own experience, and help shape how tasks and schedules are organized. This kind of collaboration would ensure the technology supports and does not replace human judgment.

7.3 Broader Implications:

The idea of a Psychosocial Digital Twin shifts the focus away from the traditional, machine-driven lens of Industry 4.0. Instead, it highlights how human factors such as cognitive ergonomics and socio-technical design can be integrated directly into the digital backbone of modern manufacturing. For Business Intelligence, this marks a real turning point: moving beyond the retrospective study of machine data toward anticipating how people will perform, adapt, and feel in the workplace. For the workforce itself, it opens the door to safer, less stressful environments that not only boost productivity but also respect natural human limits and strengths.

These findings also contribute to the growing field of **affective and physiological computing**, which aims to design systems that sense and adapt to human emotional and cognitive states (Fairclough, 2022). By integrating these insights with **trust-based adaptive automation frameworks** (Hancock, Billings, & De Visser, 2021); (Parasuraman & Cosenzo, 2020), the PDT framework extends human–AI collaboration toward a new generation of **empathetic, human-aware industrial systems**. As work environments increasingly intertwine cognitive, emotional, and algorithmic dimensions, frameworks such as PDT represent a step toward **ethical and empathetic Industry 5.0 systems**.

Disclaimer:

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of their affiliated institutions or organizations. The authors declare that there are no conflicts of interest related to this work.

References

Aricò, P., Borghini, G., & Di Flumeri, G. (2023). Passive brain-computer interfaces for mental workload and vigilance assessment: A systematic review. *Sensors*, 23:3115. 10.3390/s23063115.

Boffet, A., Arsac, L. M., & Ibanez, V. (2025). Detection of Cognitive Load Modulation by EDA and HRV. *Sensors (Basel, Switzerland)*, 25:2343. 10.3390/s25082343.

Borth, M., Gkion, M., & Weidner, R. (2022). A digital twin for real-time ergonomic analysis in human-robot collaboration. *Procedia CIRP*, 107:812-817. 10.1016/j.procir.2022.05.035.

Cao, B., Liu, Z., Han, X., Zhou, S., Zhang, H., & Wang., H. (2024). Foresee and Act Ahead: Task Prediction and Pre-Scheduling Enabled Efficient Robotic Warehousing. *Arxiv.org*.

Capulli, E., Druda, Y., & Palmese, F. (2025). Ethical and Legal Implications of Health Monitoring Wearable Devices: A Scoping Review. *Social Science & Medicine*, 370:117685. 10.1016/j.socscimed.2025.117685.

Di Stasi, L. L., Diaz-Piedra, C., & Rieiro, H. (2021). Eye-tracking metrics as indicators of mental fatigue: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 127:629-647. 10.1016/j.neubiorev.2021.05.008.

Diener, E., Wirtz, D., & Tov, W. (2009). New measures of well-being: Flourishing and positive and negative feelings. *Social Indicators Research*, 39:247-266. 10.1007/s11205-009-9582-y.

Emmanouilidis, C., Montini, E., & Cutrona, V. (2024). Manufacturing workers fatigue: an exploratory study on predictive machine learning and cross-subject generalization with implications for work design. *IFAC-PapersOnLine*, 58:557-562. 10.1016/j.ifacol.2024.09.271.

Fairclough, S. H. (2022). Affective and physiological computing for human-computer interaction. *Computers in Human Behavior*, 135:107386. 10.1016/j.chb.2022.107386.

Grandjean, E. (1973). Fatigue in industry. *British Journal of Industrial Medicine*, 36:175-186. 10.1136/oem.36.3.175.

Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. *Kahlen, F.-J. (ed): Springer*, 85-113. 10.1007/978-3-319-38756-7_4.

Hancock, P. A., Billings, D. R., & De Visser, E. J. (2021). Human-automation trust and workload: A meta-analysis. *Computers in Human Behavior*, 119:106716. 10.1016/j.chb.2021.106716.

Ivanov, D., & Dolgui, A. (2021). A digital supply chain twin for managing disruption risks and resilience in the era of Industry 4.0. *Production Planning & Control*, 32:775-788. 10.1080/09537287.2020.1768450.

Leng, J., Wang, D., & Shen, W. (2021). Digital-twins-based smart manufacturing system design in Industry 4.0: A review. *Journal of Manufacturing Systems*, 60:119-137. 10.1016/j.jmsy.2021.05.011.

Nataliya, N., & Tan, J. (2024). No Simple Fix. *How AI Harms Reflect Power and Jurisdiction in the Workplace*, 422-432. 10.1145/3630106.3658915.

Parasuraman, R., & Cosenzo, K. A. (2020). Adaptive human-automation interaction. *Frontiers in Psychology*, 11:2130. 10.3389/fpsyg.2020.02130.

Strohmeier, S. (2020). Smart HRM: A conceptual framework for human resource management in the digital age. *Human Resource Management Review*, 30:100707. 10.1016/j.hrmr.2019.100707.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12:257-285. 10.1207/s15516709cog1202_4.

What is cognitive overload, and how can we avoid it? (2021). https://blog.simplitaught.com/what-is-cognitive-overload/.

Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal mining. *Human Relations*, 4:3-38. 10.1177/001872675100400101.