BRITISH RESEARCH REVIEW

British Research Review - Issue 2026

Bio-Cognitive Integration in Virtual Conflict Simulations: A Theoretical Framework for AI-Driven Behavioural Training

E. R. Ginige* and S. Pathirane

London Institute of Business & Technology

*Corresponding author: E. R. Ginige, London Institute of Business & Technology, Hillary House, Prospect Hill, Douglas, IM1 1EQ Isle of Man. Email: eranda@libt.co.uk. ORCID: 0009-0008-2786-1886

2026

ABSTRACT

Conflict management training frequently fails to produce lasting behavioural change due to the transfer gap: theoretical knowledge proves ineffective when sympathetic nervous system activation during confrontation diminishes prefrontal cortex function in favour of amygdala-driven reactivity. This paper proposes a theoretical framework for an AI-driven, voice-activated conflict simulation integrating Polyvagal Theory, the Thomas-Kilmann Conflict Mode Instrument, and Principled Negotiation. The framework incorporates real-time biofeedback via consumer-grade wearables, with dynamic cultural adaptation informed by Meyer's Culture Map. We introduce Synthetic Experiential Learning (SEL), distinguished from existing approaches through systematic integration of physiological monitoring with adaptive AI-generated scenarios. The paper articulates theoretical foundations, proposes technical architecture, addresses limitations and contested assumptions, and outlines accelerated validation methodology incorporating AI-assisted analysis. We present this framework to invite critical peer review prior to prototype development, acknowledging the empirical work required to test its core hypotheses.

Keywords: conflict resolution, synthetic experiential learning, biofeedback, polyvagal theory, transfer of learning

1. INTRODUCTION

Workplace conflict imposes substantial costs on organisations through lost productivity, employee turnover, and diminished engagement (Bollen et al., 2016; De Dreu & Gelfand, 2008). The 2024 TalentLMS and Vyond research on What employees want from L&D reports that 41% of employees identified deficits in social and interpersonal skills within their organisations. Established models including Crucial Conversations (Patterson et al., 2021), Nonviolent Communication (Rosenberg, 2015), and Principled Negotiation (Fisher & Ury, 1981) offer sophisticated frameworks for managing disagreement. However, evidence suggests these frameworks frequently fail under the physiological stress conditions that characterise authentic interpersonal conflict.

The mechanism underlying this failure is documented in neuroscience. When an individual perceives interpersonal threat, the amygdala initiates a sympathetic nervous system response that suppresses prefrontal cortex function, the brain region responsible for executive reasoning and emotional regulation (Arnsten, 2009; Phelps & LeDoux, 2005). This means trained professionals may revert to reactive communication patterns precisely when their training is most needed. The implication is that conflict management is not primarily a cognitive challenge; it is a psychophysiological one.

This insight exposes a limitation in conventional training design. Classroom instruction and role-play exercises operate within what we term a low-stress transfer environment: participants learn techniques in conditions of physiological safety, then must deploy them under conditions of autonomic arousal and genuine interpersonal risk. Research on training transfer demonstrates that the gap between learning and application environments predicts transfer failure (Baldwin & Ford, 1988; Blume et al., 2010).

Recent developments in three fields present an opportunity to reconceptualise conflict training. First, Polyvagal Theory (Porges, 2011, 2023) provides a framework for understanding how autonomic state governs social engagement capacity. Second, consumer-grade wearable technology enables monitoring of physiological indicators including heart rate variability outside clinical settings (Goessl et al., 2017; Lalanza et al., 2023). Third, large language models demonstrate capacity to generate contextually responsive dialogue simulating human social behaviour (Croissant & Frister, 2024; Schlegel et al., 2025).

This paper proposes a Bio-Cognitive Integration (BCI) framework synthesising these developments. The core proposition is that effective conflict resolution training must simultaneously engage three systems: the biological (autonomic regulation), the strategic (conflict mode selection), and the linguistic (communicative competence under stress). By creating a simulation environment that is physiologically activating yet physically safe, we aim to enable Synthetic Experiential Learning: the formation of embodied competencies through technologically mediated experience producing transfer-ready behavioural patterns.

We present this framework as a theoretical proposal to invite critical peer review. No prototype currently exists. Our objective is to establish theoretical coherence and identify potential weaknesses before undertaking empirical development. Throughout, we acknowledge contested assumptions and the substantial validation required to test our hypotheses.

2. THEORETICAL FOUNDATIONS

The proposed framework operates across three processing layers, each grounded in established theory. This section examines each layer, articulates the rationale for integration, and addresses limitations and contested claims.

2.1 The Biological Layer: Polyvagal Theory and Biofeedback

Polyvagal Theory (PVT), developed by Porges (1995, 2011, 2023), proposes an evolutionarily informed hierarchy of three autonomic circuits governing responses to safety and threat. The ventral vagal complex facilitates social engagement and calm interaction. The sympathetic nervous system activates fight-or-flight responses under perceived threat. The dorsal vagal complex triggers immobilisation under overwhelming danger.

Central to PVT is neuroception: the process by which the nervous system evaluates environmental cues for safety or danger without conscious appraisal (Porges, 2011, 2025). In workplace conflict, neuroception can shift an individual from a ventral vagal state, where empathic listening and problem-solving are possible, to a sympathetic state characterised by defensive or aggressive responses, irrespective of cognitive intentions.

2.1.1 Critical Evaluation of Polyvagal Theory

PVT has attracted scientific critique requiring transparent acknowledgement. Grossman (2023) challenged foundational premises regarding the anatomical specificity of ventral versus dorsal vagal pathways and the use of respiratory sinus arrhythmia (RSA) as a sole index of vagal tone. Taylor et al. (2022) questioned whether the phylogenetic hierarchy PVT describes accurately reflects autonomic nervous system evolution across species.

Porges (2025) has responded by introducing refined metrics including weighted coherence and vagal efficiency, and by clarifying the distinction between RSA as an empirical bridge and as a theoretical component. Importantly, PVT's predictive utility for intervention design may be independent of whether its evolutionary narrative is fully accurate. As Giroux et al. (2023) observe, PVT may function as a useful heuristic model even where specific anatomical claims remain contested.

For this framework, we adopt PVT with the following qualifications. We treat ventral vagal, sympathetic, and dorsal vagal states as observable physiological patterns indexed by validated measures (HRV, electrodermal activity) rather than as claims about neuroanatomical pathways. We acknowledge that the framework's validity does not depend on PVT's evolutionary narrative being correct; it depends on whether the autonomic state distinctions PVT describes can be reliably detected and whether training that targets these states improves conflict resolution outcomes. These are empirical questions the validation programme will address.

2.1.2 Biofeedback Integration

The practical application utilises consumer-grade wearable devices measuring HRV and electrodermal activity (EDA). Ferreira et al. (2025) systematically reviewed biofeedback-based workplace interventions, finding HRV biofeedback demonstrates significant stress reduction effects. Goessl et al. (2017) reported large effect sizes for HRV biofeedback training on self-reported stress and anxiety (Hedges' g = 0.83).

In the proposed system, physiological signals serve dual functions: as real-time input to adaptive difficulty algorithms, and as longitudinal measures of developing autonomic regulation capacity. When physiological indicators suggest sympathetic activation (elevated heart rate, reduced HRV, increased EDA), the system can respond in two calibrated modes. In Co-regulation Mode, the AI adjusts vocal prosody, content, and pace to reduce affective challenge, modelling interpersonal co-regulation (Dana, 2018). In Challenge Mode, the AI maintains or escalates difficulty to develop the user's capacity to sustain regulated function under pressure.

Several technical limitations require acknowledgement. Consumer-grade devices are subject to motion artefacts affecting signal quality, and the temporal resolution required for natural conversational adaptation may exceed current capabilities. Baseline physiological patterns vary substantially across individuals, requiring personalised calibration. These challenges may prove more difficult than current literature suggests.

2.2 The Strategic Layer: Conflict Mode Identification

The Thomas-Kilmann Conflict Mode Instrument (TKI) identifies five behavioural orientations mapped along two dimensions: assertiveness and cooperativeness (Thomas & Kilmann, 1974). These modes are: Competing (high assertiveness, low cooperativeness), Collaborating (high on both), Compromising (moderate on both), Avoiding (low on both), and Accommodating (low assertiveness, high cooperativeness).

Within the BCI framework, the AI's behaviour is parametrically mapped to these five modes. The user's task is twofold: correctly identify the AI's operative mode through observation, and select an appropriate strategic response. Importantly, collaboration is not universally optimal. The system rewards strategic flexibility, the capacity to match mode to situation, rather than default reliance on any single approach.

This layer integrates with Principled Negotiation (Fisher & Ury, 1981; Fisher et al., 2011). The AI evaluates whether user proposals address underlying interests rather than positional demands, generate options for mutual gain, and reference objective criteria. Susskind and Dinnar (2024) at MIT demonstrated the viability of using generative AI for negotiation coaching, finding that AI preparation tools enhanced student preparedness.

A limitation requiring acknowledgement is that TKI modes represent analytical categories imposed on continuous behavioural variation. Individuals rarely exhibit pure modes; behaviour is typically mixed and context-dependent. The same verbal content may reflect different underlying orientations depending on tone and relationship. The framework must therefore treat mode classification as probabilistic rather than categorical, with the AI inferring likely mode from multiple behavioural indicators whilst acknowledging uncertainty.

2.3 The Linguistic Layer: Communicative Competence Under Stress

The third layer evaluates verbal output using natural language processing (NLP) calibrated against communication frameworks:

2.3.1 Observations versus Evaluations

Rosenberg's (2015) Nonviolent Communication distinguishes objective observations from evaluative judgements. The NLP engine flags evaluative language and tracks the user's capacity to maintain observational framing under affective pressure.

2.3.2 Psychological Safety Signals

Drawing on Edmondson's (1999, 2019) research, and the comprehensive review synthesising 185 papers on psychological safety (Edmondson & Bransby, 2023), the system evaluates whether language creates conditions reducing perceived interpersonal risk. This includes acknowledgement of the other party's perspective, invitation for input, and tolerance of disagreement.

2.3.3 Direct Challenge with Personal Care

Scott's (2017) Radical Candor framework identifies four quadrants defined by caring personally and challenging directly. The NLP engine classifies output along these dimensions, tracking movement between Radical Candor, Ruinous Empathy, Obnoxious Aggression, and Manipulative Insincerity.

Research supports the feasibility of this approach. Schlegel et al. (2025) found contemporary LLMs achieved 81% accuracy on standardised emotional intelligence tests compared to 56% for human participants. Llanes-Jurado et al. (2024) developed conversational virtual humans integrating personality and emotional expression, producing interactions validated by participant assessment. These findings suggest technical capacity for nuanced emotional simulation is approaching practical utility.

A limitation concerns the cultural specificity of these frameworks. NVC's distinction between observations and evaluations may be more ambiguous in high-context cultures where indirect communication is normative. Radical Candor's emphasis on direct challenge may conflict with communication norms emphasising face-saving. This motivates the cultural adaptation mechanisms discussed in Section 4.

2.4 Synthetic Experiential Learning: Theoretical Justification

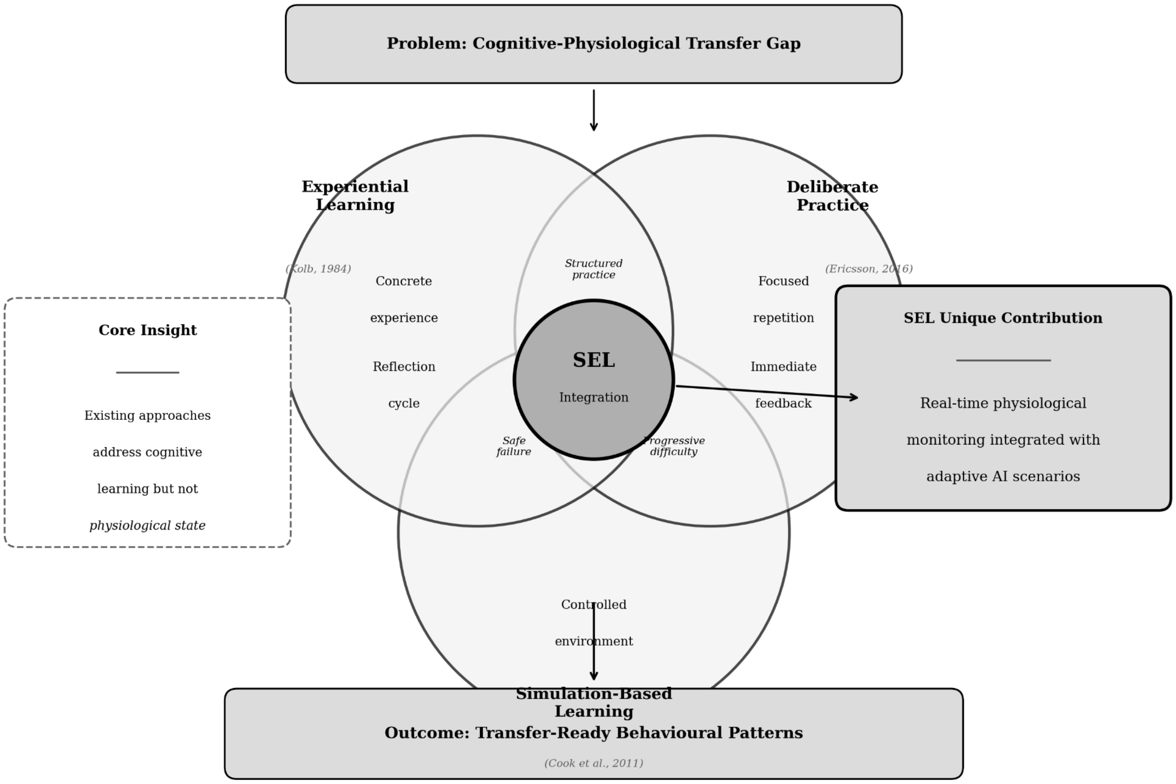
We introduce the term **Synthetic Experiential Learning** (SEL) to describe the framework's pedagogical approach. This section justifies why SEL merits distinction from existing constructs including experiential learning (Kolb, 1984), deliberate practice (Ericsson & Pool, 2016), and simulation-based learning.

2.4.1 Distinguishing Features of Synthetic Experiential Learning

Kolb's (1984) experiential learning theory relies on naturally occurring or structured experiences, whose quality varies with facilitator skill and situational factors. Ericsson's deliberate practice framework emphasises focused repetition with immediate feedback (Ericsson & Pool, 2016), but generating consistent practice opportunities in interpersonal domains is difficult. Simulation-based learning produces better outcomes than traditional instruction (Cook et al., 2011), yet existing approaches focus on procedural skills with objective criteria and rarely incorporate physiological monitoring.

SEL synthesises elements from each approach whilst introducing a novel integration. From experiential learning, SEL takes the emphasis on concrete experience as the foundation for learning. From deliberate practice, SEL adopts the principles of focused repetition, immediate feedback, and progressive difficulty. From simulation-based learning, SEL inherits the controlled environment enabling safe failure and iteration.

The distinctive contribution of SEL is the systematic integration of physiological monitoring with adaptive AI-generated scenarios. This integration addresses a specific gap: the disconnect between cognitive learning and physiological performance. By monitoring autonomic state during practice, SEL can ensure that users develop competence not merely in what to say, but in maintaining the physiological regulation necessary to execute communication strategies under stress. The synthetic environment generates experiences that are artificial in origin but authentic in physiological activation.



SEL = Synthetic Experiential Learning

Figure 1. Synthetic Experiential Learning: Theoretical Positioning

2.4.2 Theoretical Grounding

SEL draws theoretical support from embodied cognition research demonstrating that cognitive processes are grounded in bodily states (Barsalou, 2008; Wilson, 2002). If conflict resolution competence requires specific physiological states for effective deployment, then training must address both cognitive content and physiological capacity. Training that develops declarative knowledge without corresponding physiological adaptation may produce competence that cannot be accessed under stress.

This aligns with research on state-dependent learning, which demonstrates that information encoded in one physiological state is more accessible when that state is reinstated (Godden & Baddeley, 1975). If conflict resolution skills are encoded primarily in calm, safe training environments, they may be less accessible when needed in activated, threatened states. SEL's design, training under controlled physiological activation, may facilitate encoding that transfers more readily to authentic conflict situations.

We acknowledge this theoretical justification requires empirical validation. The hypothesis that physiologically activated training produces superior transfer to authentic conflict situations, compared with cognitively focused training in calm conditions, is testable and forms a core component of the validation programme.

2.5 Cultural Adaptation Framework

Conflict norms, communication styles, and appropriate resolution strategies vary substantially across cultures. A framework claiming general applicability must incorporate systematic cultural adaptation. This section reviews relevant cultural frameworks and articulates the adaptation mechanisms integrated throughout the BCI architecture.

2.5.1 Review of Cultural Frameworks

Three frameworks offer empirically grounded dimensions for understanding cultural variation in communication and conflict:

- **Hofstede's Cultural Dimensions:** Hofstede's (1980, 2001) research identified dimensions including Power Distance and Individualism-Collectivism, providing country-level scores for quantitative comparison. Limitations include reliance on dated data and ecological fallacy risks (McSweeney, 2002).
- **GLOBE Study:** The GLOBE study (House et al., 2004) extended Hofstede's work across 62 societies, distinguishing cultural practices from values and identifying dimensions relevant to conflict including Assertiveness and Institutional Collectivism. It offers stronger methodological grounding but shares limitations regarding individual variation.
- **Meyer's Culture Map:** Meyer's (2014) framework identifies eight scales particularly relevant to workplace communication: Communicating (low-context to high-context), Evaluating (direct to indirect negative feedback), Persuading (principles-first to applications-first), Leading (egalitarian to hierarchical), Deciding (consensual to top-down), Trusting (task-based to relationship-based), Disagreeing (confrontational to avoids-confrontation), and Scheduling (linear-time to flexible-time).

2.5.2 Framework Selection: Meyer's Culture Map

For the BCI framework, we select Meyer's Culture Map as the primary cultural adaptation mechanism for several reasons. First, its dimensions map directly to observable communication behaviours relevant to conflict scenarios (directness of feedback, comfort with open disagreement, communication context). Second, its relative recency and practitioner orientation provide accessible implementation guidance. Third, its eight scales offer sufficient granularity for meaningful adaptation without excessive complexity.

We acknowledge Meyer's framework has received less academic validation than Hofstede or GLOBE. Country placements on Meyer's scales derive from Meyer's consulting experience and targeted research rather than large-scale psychometric studies. This limitation is mitigated by the framework's face validity and alignment with more extensively validated constructs from cross-cultural psychology.

2.6 Integration Throughout the Architecture

Cultural adaptation operates across all three processing layers:

1. **Biological Layer Adaptation:** Baseline physiological patterns may vary across populations due to both genetic and cultural factors affecting autonomic function. The system establishes individualised baselines rather than assuming universal norms. Additionally, what constitutes appropriate arousal during conflict may differ; cultures with confrontational disagreement norms may show higher baseline activation during disagreement without this indicating dysregulation.
2. **Strategic Layer Adaptation:** The appropriateness of TKI modes varies by cultural context. Cultures high on Meyer's 'avoids confrontation' scale may find Competing strategies relationally damaging even when situationally warranted. Cultures emphasising hierarchical leadership may require different strategic approaches when conflict involves status differentials. The AI's evaluation of strategic appropriateness incorporates cultural weighting.
3. **Linguistic Layer Adaptation:** The NLP engine's evaluation criteria adjust based on cultural parameters. In high-context cultures, indirect communication that might score poorly on directness metrics may represent culturally appropriate conflict management. The distinction between observations and evaluations in NVC may require recalibration for cultures where embedded evaluation is communicatively normative. Psychological safety signals differ across cultures; explicit invitation for disagreement may be unnecessary or even confusing in cultures with strong disagreement norms.

Implementation requires user cultural profile specification at system initialisation. Users indicate their own cultural background and, where relevant, the cultural context of typical conflict partners. This enables scenario generation matching the user's actual intercultural conflict experiences. An important design principle is that the system trains users for their real-world contexts rather than imposing any single cultural norm as universally correct.

2.7. Proposed Technical Architecture

This section describes the proposed technical architecture, comparative alternatives, and acknowledged technical challenges requiring resolution during prototype development.

2.7.1 Comparative Analysis of Simulation Approaches

Three principal approaches exist for AI-driven conflict simulation, each with distinct trade-offs:

1. **Branching Narrative (Heuristic):** Pre-authored dialogue trees offering pedagogical control but limited ecological validity and replayability. Established in medical simulation (Sanchez et al., 2023) but insufficient for conflict training's nuanced variation.
2. **Generative LLM (Dynamic):** Real-time generation offering ecological validity and variation, but risking inconsistent responses and difficulty constraining behaviour to pedagogical parameters. Research supports feasibility for negotiation training (Susskind & Dinnar, 2024; Croissant & Frister, 2024).
3. **Hybrid Affective-Loop (Proposed):** LLM generates dialogue whilst a Bio-Controller layer constrains emotional state and maintains pedagogical structure. This approach balances realism with pedagogical coherence but introduces technical complexity requiring real-time synchronisation between biofeedback data and text generation.

2.7.2 The Hybrid Affective-Loop Architecture

The proposed architecture comprises four integrated components:

1. **Biofeedback Acquisition Layer:** Consumer-grade wearables transmit HRV and EDA data via Bluetooth to the processing system. Signal processing algorithms filter motion artefacts and extract relevant metrics. Technical specifications include minimum 4Hz sampling for HRV and sub-second latency for state classification.
2. **State Classification Engine:** A machine learning classifier, trained on labelled physiological data, categorises user autonomic state (ventral vagal, sympathetic activation, dorsal vagal indicators). Classification confidence scores inform downstream processing. Individual calibration addresses between-person variability in baseline physiology.
3. **Bio-Controller Layer:** This layer translates physiological state into constraints for dialogue generation. In Co-regulation Mode, it instructs the LLM to reduce affective intensity, slow pace, and increase warmth. In Challenge Mode, it instructs the LLM to maintain or increase difficulty within defined pedagogical bounds. The Bio-Controller also maintains fidelity to the assigned TKI mode and cultural parameters.
4. **Generative Dialogue Engine:** A large language model generates contextually appropriate dialogue within constraints specified by the Bio-Controller. Prompt engineering ensures the AI maintains consistent persona, emotional tone, and strategic orientation whilst generating varied surface expressions.

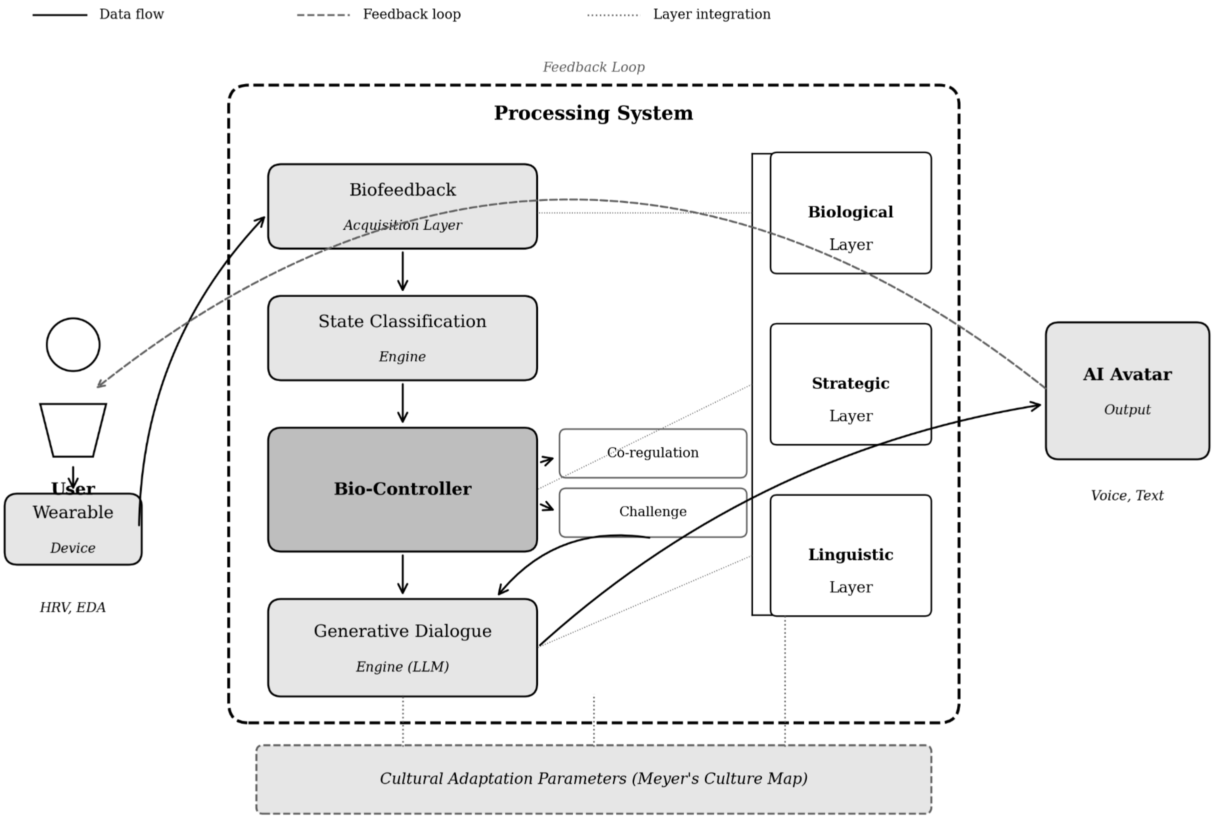


Figure 2. Bio-Cognitive Integration Architecture

2.7.3 Technical Challenges

Several technical challenges require resolution during prototype development:

1. **Latency Requirements:** Natural conversation requires response latencies below approximately 500ms. Biofeedback processing, state classification, Bio-Controller decision-making, and LLM generation must be completed within this window. Current LLM inference speeds approach but may not consistently achieve this target.
2. **Signal Quality from Consumer Devices:** Clinical-grade physiological monitoring equipment substantially outperforms consumer wearables. Whether consumer devices provide sufficient signal quality for reliable state classification in dynamic, conversational contexts remains an empirical question. Lalanza et al. (2023) provide methodological guidance for HRV biofeedback measurement, but their recommendations primarily address clinical-grade equipment.
3. **LLM Consistency and Constraint Adherence:** Ensuring LLMs maintain consistent emotional presentation, TKI mode adherence, and cultural appropriateness across

extended interactions is challenging. Prompt engineering and fine-tuning approaches show promise but may not achieve the consistency required for pedagogically coherent training.

We explicitly acknowledge these challenges may prove more difficult than anticipated. The validation methodology includes technical feasibility assessment as a gating criterion before full-scale pedagogical evaluation.

2.8. Framework for Longitudinal Competency Development

Conventional conflict training suffers from its episodic nature. Participants acquire declarative knowledge and are expected to maintain skills independently. Research demonstrates one-off interventions produce declining returns without sustained reinforcement (Blume et al., 2010; Kauffeld & Lehmann-Willenbrock, 2024). The BCI framework addresses this through a Cumulative Competency Model tracking development across sessions and progressively adjusting difficulty.

2.8.1 Competency Dimensions

The system tracks three parallel competency dimensions:

- **Autonomic Regulation Competency:** The user's capacity to maintain or rapidly return to a ventral vagal state under progressively escalating affective challenge. Baseline HRV profiles are established initially, and improvement is tracked longitudinally. Metrics include time to recover baseline HRV following simulated conflict, variability of physiological state during conversation, and the maximum challenge level at which regulated function is maintained.
- **Strategic Flexibility Competency:** The accuracy and speed of conflict mode identification, appropriateness of strategic response selection across varied scenarios, and capacity to shift strategy mid-interaction in response to changed conditions. Assessment includes recognition accuracy for AI-presented modes and evaluation of strategic fit between user responses and scenario demands.
- **Communicative Competency:** NLP analysis of verbal output across NVC, Psychological Safety, and Radical Candor dimensions, with emphasis on performance under elevated physiological arousal. The key metric is the gap between communication quality in calm versus activated states; competency development aims to narrow this gap.

2.8.2 Statistical Modelling of Competency Transitions

We propose a Hidden Markov Model (HMM) to track user transitions between competency states across sessions. HMMs are appropriate for this application because they model systems where underlying states (true competency levels) are not directly observable but can be inferred from observable outputs (physiological data, verbal responses, strategic choices).

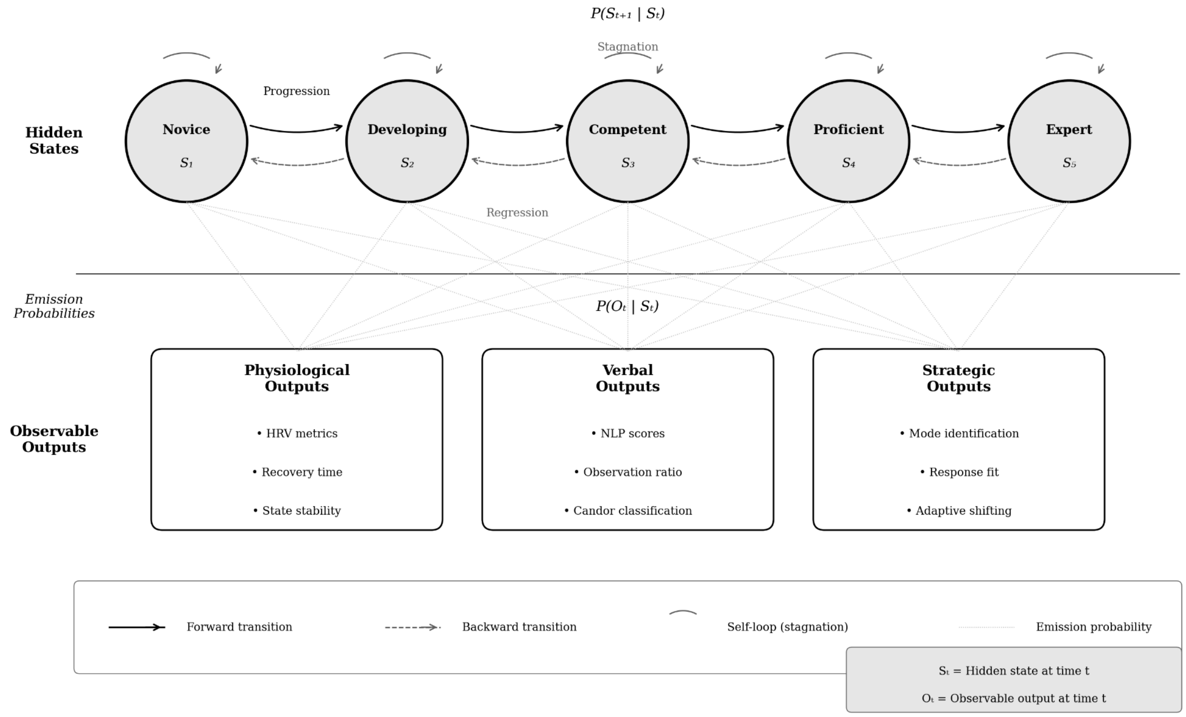


Figure 3. Hidden Markov Model for Competency Tracking

- **State Space Definition:** The hidden state space comprises a unified competency construct with five ordinal levels: Novice, Developing, Competent, Proficient, and Expert. This unified approach, rather than separate tracking of each competency dimension, reflects the theoretical integration of the three layers. Physiological regulation, strategic flexibility, and communicative competence are hypothesised to develop interdependently; users who cannot regulate physiologically will struggle to deploy sophisticated strategies or maintain linguistic precision.
- **Observable Outputs:** Three categories of observable outputs inform state inference. Physiological outputs include HRV metrics during and following scenarios, time to physiological recovery, and state stability across sessions. Verbal outputs include NLP-derived scores for observational language, psychological safety signals, and candor quadrant classification. Strategic outputs include TKI mode identification accuracy, response appropriateness scores, and adaptive strategy shifting during scenarios.
- **Transition Probability Estimation:** Initial transition probabilities will be specified based on skill acquisition research (progression is more likely from adjacent states; regression is less common than stagnation), with Bayesian updating as data accumulates.

- **Emission Probability Specification:** Emission probabilities define the likelihood of observing particular outputs given competency state. Initial values will be derived from expert judgement and existing literature, then refined through empirical calibration.
- **Model Application:** The HMM serves three functions. First, it provides posterior probability estimates of the user's current competency state given all observed data, informing adaptive difficulty calibration. Second, it predicts likely development trajectories, enabling personalised training path recommendations. Third, it identifies users whose transition patterns deviate from expected progression, flagging potential issues with the training programme or individual learning barriers.

We acknowledge this model specification is entirely theoretical at present. The validation programme must establish whether the hypothesised state structure is empirically supported, whether the three output categories provide sufficient information for reliable state inference, and whether the transition dynamics assumed align with actual user development patterns. Model comparison approaches (e.g., comparing five-state versus three-state alternatives) will be employed to identify the most parsimonious structure supported by data.

3. METHODOLOGY

Before deployment, the framework requires empirical validation. We propose an accelerated three-phase methodology incorporating AI-assisted analysis to reduce time and cost whilst maintaining rigour.

3.1 Phase I: Technical Feasibility and Expert Calibration (Months 1-6)

This phase establishes whether the proposed technical architecture is achievable and calibrates system parameters using expert judgement.

Technical Feasibility Assessment	Prototype development will address critical technical questions: Can consumer wearables provide sufficient signal quality for reliable state classification? Can the system achieve sub-500ms response latency? Can the LLM maintain consistent TKI mode presentation across extended interactions? This assessment constitutes a gating criterion; if technical feasibility cannot be established, subsequent phases will not proceed as planned.
Expert Calibration Protocol	Subject matter experts in conflict mediation, negotiation, and cross-cultural communication will interact with the prototype to calibrate scoring parameters. Expert selection criteria include minimum five years of professional practice and demonstrated expertise through publication, certification, or recognised professional standing. We will recruit 20-25 experts representing diverse cultural backgrounds to ensure calibration is not culturally biased.
Calibration Objectives	Experts will evaluate and provide feedback on NLP engine scoring accuracy for communicative competence constructs, AI avatar TKI mode recognisability and behavioural accuracy, cultural adaptation appropriateness across represented cultural contexts, and the relationship between biofeedback parameters and perceived scenario difficulty. Inter-rater reliability will be assessed using intraclass correlation coefficients; minimum acceptable ICC = 0.70.
AI-Assisted Analysis	Expert feedback sessions will be transcribed and analysed using LLM-based qualitative analysis to identify themes, areas of expert agreement and disagreement, and specific calibration recommendations. This approach accelerates analysis whilst maintaining systematic rigour. Human researchers will verify AI-generated analyses against source data.

3.2 Phase II: Controlled Efficacy Trial (Months 7-14)

This phase tests whether the BCI framework produces superior learning outcomes compared with established alternatives.

Study Design	A randomised controlled trial with three arms: (a) BCI Simulation training, (b) traditional classroom-based conflict management training (active control), and (c) video-based self-directed learning (active control). Three-arm design enables comparison with both high-contact and low-contact alternatives.
Sample Size Determination	Power analysis assuming medium effect size (Cohen's d = 0.5), alpha = 0.05, power = 0.80, and three-group comparison indicates minimum n = 52 per arm (total N = 156). To account for attrition (estimated 20%), we will recruit N = 195 (65 per arm). If Phase I suggests effect sizes differ from assumptions, power analysis will be recalculated.
Outcome Measures	Primary outcomes include Post-Intervention Stress Recovery Rate, Situational Judgement Test performance, and Behavioural Observation Ratings from trained raters evaluating role-play interactions. Secondary outcomes include conflict self-efficacy, TKI mode identification accuracy, NLP-derived competency scores, and participant satisfaction.
Analysis Plan	Mixed-effects models will account for repeated measures and individual variation. Intention-to-treat analysis will be primary; per-protocol analysis will be conducted as sensitivity check. Pre-registration of hypotheses and analysis plan on Open Science Framework before data collection.

3.3 Phase III: Field Transfer Validation (Months 15-22)

This phase tests whether improvements observed in controlled conditions transfer to authentic workplace interactions.

Design	Pragmatic field trial in two organisational contexts selected for high conflict salience: healthcare management teams (where interpersonal conflict affects patient safety and staff wellbeing) and client-facing professional services teams (where conflict management is core competency). Participants who completed Phase II BCI training will be compared with matched controls from the same organisations who did not participate.
Outcome Measures	360-degree feedback surveys at baseline, three months, and six months focusing on conflict-relevant behaviours. Team psychological safety surveys using Edmondson's validated instrument. Organisational metrics including conflict incident reports, staff turnover, and (for healthcare) patient safety indicators. Client or patient satisfaction data where available.
Limitations	Field validation lacks the control of laboratory conditions. Confounding factors including organisational changes, personnel turnover, and external stressors may affect outcomes. Results will be interpreted with appropriate caution regarding causal inference. Qualitative interviews with participants and supervisors will supplement quantitative measures to understand mechanisms and contextual factors.

Iterative Refinement	Throughout all phases, findings will inform framework refinement. We anticipate the need for substantial modification based on empirical feedback. The validation programme is designed to test hypotheses, not to confirm predetermined conclusions.

3.4 Ethical Considerations and Limitations

3.4.1 Data Protection and Privacy

The collection and processing of physiological data raises significant privacy concerns. All data processing will comply with relevant data protection regulations including GDPR and equivalent frameworks. Specific protections include explicit informed consent with clear explanation of data types collected, purpose limitation, and retention periods; local processing of physiological data where technically feasible to minimise transmission of sensitive information; data minimisation principles with collection of only data necessary for training functions; and user control including ability to delete personal data and opt out of longitudinal tracking.

Physiological data may inadvertently reveal health conditions. Users must be informed of this possibility during consent processes. The system will not provide health diagnoses or medical advice based on physiological data.

3.4.2 User Psychological Safety

The framework intentionally induces managed physiological stress as a training mechanism. This creates duty-of-care obligations. The system must include safeguards against excessive or prolonged stress activation. If physiological indicators suggest user distress exceeding training benefit, the system should reduce difficulty or pause the scenario. Users should have immediate access to exit any scenario without penalty.

Importantly, the system must avoid pathologising normal stress responses. Not all sympathetic activation represents failure; healthy conflict engagement involves tolerable arousal. Calibration must distinguish productive challenge (within the user's window of tolerance) from counterproductive distress, avoiding implications that emotional arousal is inherently problematic.

Persistent sympathetic dysregulation detected during training may indicate mental health concerns unrelated to the training programme. Protocols must address appropriate response, which may include pausing training and recommending professional support, whilst avoiding diagnostic overreach.

3.4.3 Cultural and Power Considerations

Chen (2025) cautions that AI-mediated training tools may oversimplify cultural, ethical, and power dimensions of interpersonal conflict. Several risks require attention.

Cultural frameworks including Meyer's Culture Map provide useful heuristics but inevitably simplify cultural complexity. Treating cultural profiles as fixed categories risks stereotyping and may not capture within-culture variation or individuals' multicultural identities. The system should present cultural adaptation as adjustable parameters rather than deterministic profiles.

Power dynamics in conflict are not merely communicative; they reflect structural inequalities that AI simulation cannot address. Training someone to communicate more effectively in conflict with a more powerful party does not address the underlying power imbalance. Users should understand the framework's scope and limitations.

3.4.4 Theoretical Limitations

Polyvagal Theory, whilst providing useful heuristics, remains contested in its anatomical specifics (Grossman, 2023). The framework's validity ultimately depends on empirical demonstration that autonomic state monitoring improves training outcomes, regardless of whether PVT's evolutionary narrative is correct.

The Thomas-Kilmann Instrument has been critiqued for static categorisation of dynamic behavioural processes. The framework's treatment of conflict modes as probabilistically inferred rather than categorically classified partially addresses this limitation.

NVC and Radical Candor frameworks have received less rigorous empirical validation than their popularity might suggest. Their inclusion reflects practitioner adoption rather than definitive evidence of efficacy.

These limitations reinforce the importance of the empirical validation programme and iterative refinement based on observed outcomes rather than theoretical assumptions.

5. CONCLUSION

This paper has proposed a Bio-Cognitive Integration framework addressing a persistent gap in conflict management training: the disconnection between cognitive learning environments and the psychophysiological conditions of authentic conflict. By integrating Polyvagal Theory, the Thomas-Kilmann Conflict Mode Instrument, and Principled Negotiation within an AI-driven simulation incorporating real-time biofeedback and cultural adaptation, the framework aims to develop what we term Synthetic Experiential Learning: embodied competencies formed through technologically mediated experience producing transfer-ready behavioural patterns.

We have been explicit about limitations and contested assumptions throughout. Polyvagal Theory provides useful heuristics but faces substantive scientific critique. Technical feasibility of the proposed architecture remains undemonstrated. The Hidden Markov Model specification is entirely theoretical pending empirical calibration. These acknowledgements are not weaknesses to be minimised; they are honest representations of the framework's current status as a theoretical proposal requiring empirical validation.

The convergence of wearable biometric technology, large language models, and neurophysiological theory creates an opportunity to reconsider how professionals develop conflict resolution competencies. Whether this opportunity can be realised, and whether the BCI framework represents an effective approach, are empirical questions we invite the research community to help us answer. We present this work to solicit critical peer review before prototype development, in the belief that rigorous theoretical scrutiny strengthens the foundation for subsequent empirical work.

Declarations

Author Contributions. E.R.G. led conceptualisation, methodology, and original draft preparation. S.P. contributed to methodology development and critical revision.

Funding. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest Statement. The London Institute of Business & Technology has commercial interest in developing the proposed system. The authors declare no other conflicts of interest.

Data Availability. No datasets were generated or analysed during this study.

Acknowledgements. The authors thank Tony Reilly OBE for support and guidance.

References

Arnsten, A. F. T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. Nature Reviews Neuroscience, 10(6), 410-422.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. Personnel Psychology, 41(1), 63-105.

Barsalou, L. W. (2008). Grounded cognition. Annual Review of Psychology, 59, 617-645.

Blascovich, J., & Mendes, W. B. (2010). Social psychophysiology and embodiment. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), Handbook of social psychology (5th ed., pp. 194-227). Wiley.

Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. Journal of Management, 36(4), 1065-1105.

Bollen, K., Euwema, M., & Munduate, L. (2016). Advancing workplace mediation through integration of theory and practice. Springer.

Chen, J. A. (2025). Preserving human autonomy in AI-mediated negotiations [Master's thesis]. Massachusetts Institute of Technology.

Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., Erwin, P. J., & Hamstra, S. J. (2011). Technology-enhanced simulation for health professions education: A systematic review and meta-analysis. JAMA, 306(9), 978-988.

Croissant, M., & Frister, M. (2024). An appraisal-based chain-of-emotion architecture for affective language model game agents. PLoS ONE, 19(5), e0301033.

Dana, D. (2018). The polyvagal theory in therapy: Engaging the rhythm of regulation. W. W. Norton.

De Dreu, C. K. W., & Gelfand, M. J. (2008). The psychology of conflict and conflict management in organizations. Lawrence Erlbaum.

Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. Administrative Science Quarterly, 44(2), 350-383.

Edmondson, A. C. (2019). The fearless organization: Creating psychological safety in the workplace for learning, innovation, and growth. Wiley.

Edmondson, A. C., & Bransby, D. P. (2023). Psychological safety comes of age: Observed themes in an established literature. Annual Review of Organizational Psychology and Organizational Behavior, 10, 55-78.

Ericsson, A., & Pool, R. (2016). Peak: Secrets from the new science of expertise. Houghton Mifflin Harcourt.

Ferreira, S., Rodrigues, M. A., Magalhaes, C., Rodrigues, P. P., & Rocha, N. B. (2025). Interventions based on biofeedback systems to improve workers' psychological well-being, mental health, and safety: Systematic literature review. Journal of Medical Internet Research, 27(1), e70134.

Fisher, R., & Ury, W. (1981). Getting to yes: Negotiating agreement without giving in. Houghton Mifflin.

Fisher, R., Ury, W., & Patton, B. (2011). Getting to yes: Negotiating agreement without giving in (3rd ed.). Penguin.

Giroux, C., Ahlers, D., & Miawotoe, A. (2023). Polyvagal approaches: Scientifically questionable but useful in practice. Journal of Psychiatry Reform, 10(11).

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. British Journal of Psychology, 66(3), 325-331.

Goessl, V. C., Curtiss, J. E., & Hofmann, S. G. (2017). The effect of heart rate variability biofeedback training on stress and anxiety: A meta-analysis. Psychological Medicine, 47(15), 2578-2586.

Gratch, J., DeVault, D., Lucas, G. M., & Marsella, S. (2015). Negotiation as a challenge problem for virtual humans. In Intelligent Virtual Agents (pp. 201-215). Springer.

Grossman, P. (2023). Fundamental challenges and likely refutations of the five basic premises of the polyvagal theory. Biological Psychology, 180, 108589.

Hofstede, G. (1980). Culture's consequences: International differences in work-related values. Sage.

Hofstede, G. (2001). Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations (2nd ed.). Sage.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (Eds.). (2004). Culture, leadership, and organizations: The GLOBE study of 62 societies. Sage.

Kauffeld, S., & Lehmann-Willenbrock, N. (2024). How to maximise the impact of workplace training: A mixed-method analysis. European Journal of Work and Organizational Psychology, 33(2), 201-217.

Kolb, D. A. (1984). Experiential learning: Experience as the source of learning and development. Prentice-Hall.

Lalanza, J. F., Lorente, S., Bullich, R., Garcia, C., Losilla, J. M., & Capdevila, L. (2023). Methods for heart rate variability biofeedback (HRVB): A systematic review and guidelines. Applied Psychophysiology and Biofeedback, 48(3), 275-297.

Llanes-Jurado, J., Marchiori, E. J., Alcaniz, M., & Guixeres, J. (2024). Developing conversational virtual humans for social emotion elicitation based on large language models. Expert Systems with Applications, 249, 123573.

McSweeney, B. (2002). Hofstede's model of national cultural differences and their consequences: A triumph of faith - a failure of analysis. Human Relations, 55(1), 89-118.

Meyer, E. (2014). The culture map: Breaking through the invisible boundaries of global business. PublicAffairs.

Patterson, K., Grenny, J., McMillan, R., & Switzler, A. (2021). Crucial conversations: Tools for talking when stakes are high (3rd ed.). McGraw-Hill.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. Neuron, 48(2), 175-187.

Porges, S. W. (1995). Orienting in a defensive world: Mammalian modifications of our evolutionary heritage. Psychophysiology, 32(4), 301-318.

Porges, S. W. (2011). The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication, and self-regulation. W. W. Norton.

Porges, S. W. (2023). The vagal paradox: A polyvagal solution. Comprehensive Psychoneuroendocrinology, 16, 100200.

Porges, S. W. (2025). Polyvagal theory: Current status, clinical applications, and future directions. Clinical Neuropsychiatry, 22(3), 169-184.

Rockstroh, C., Blum, J., & Goritz, A. S. (2023). The effectiveness and user experience of a biofeedback intervention program for stress management supported by virtual reality and mobile technology: A randomized controlled study. BMC Digital Health, 1, Article 42.

Rosenberg, M. B. (2015). Nonviolent communication: A language of life (3rd ed.). PuddleDancer Press.

Sanchez, D. R., Rueda, A., Kawasaki, K., Van Lysebetten, S., & Diaz, D. (2023). Implications for workplace training. Multimodal Technologies and Interaction, 7(5), 50.

Schlegel, K., Sommer, N. R., & Mortillaro, M. (2025). Large language models are proficient in solving and creating emotional intelligence tests. Communications Psychology, 3(1), Article 80.

Scott, K. (2017). Radical candor: Be a kick-ass boss without losing your humanity. St. Martin's Press.

Susskind, L. E., & Dinnar, S. (2024). Negotiation coaching bots [Preprint]. MIT Graduate Education in the Age of AI.

TalentLMS & Vyond. (2024). What employees want from L&D in 2024.

Taylor, E. W., Leite, C. A. C., & Skovgaard, N. (2010). Autonomic control of cardiorespiratory interactions in fish, amphibians and reptiles. Brazilian Journal of Medical and Biological Research, 43(7), 600-610.

Thomas, K. W., & Kilmann, R. H. (1974). Thomas-Kilmann conflict mode instrument. Xicom.

Wilson, M. (2002). Six views of embodied cognition. Psychonomic Bulletin & Review, 9(4), 625-636.