Future of Work:

The Journal of Labor Transformation, Technology Integration, and Human Adaptation

eISSN: 3143-8148 (Online)

doi.org/10.70175/futureofworkjournal.2026

Volume 1 Issue 1 - September 2026

Using AI-Simulated Personas to Explore Motivational Interventions: A Methodological Investigation

Jonathan H. Westover, Utah Valley University, Future State University, Nexus Institute for Work & AI, Catalyst Center for Work Innovation

Received March 5, 2026
Accepted for publication July 10, 2026
Published Early Access August 3, 2026

Abstract: *Experimental research on workplace motivation faces significant practical and ethical constraints, as random assignment to interventions, manipulation of organizational contexts, and testing across diverse populations are often infeasible in field settings. This paper investigates whether AI-simulated worker personas can serve as a complementary methodological tool for early-stage exploration of motivational interventions, examining this question using beneficiary impact and job crafting interventions in a simulated fundraising context. We generated 240 diverse worker personas using three large language models (Claude, GPT-4, Llama), varying systematically in personality traits, cultural backgrounds, age, and prior experiences. Personas were randomly assigned to control training, beneficiary impact intervention, or job crafting reflection intervention conditions, with motivation, anticipated performance, and response quality measured through quantitative ratings and qualitative text analysis. Across all three LLMs, beneficiary impact and job crafting conditions showed substantially higher motivation ratings than control (Cohen's d = 3.15 and 3.97 respectively), with consistent patterns across model architectures and LLM choice explaining only 3-4% of variance. Individual differences moderated*

intervention effects in theoretically predictable ways, with collectivist personas and those high in agreeableness showing stronger responses, while qualitative analysis revealed distinct psychological mechanisms and generated testable hypotheses about intervention processes. AI-simulated personas show promise for rapid hypothesis generation and iterative exploration of motivational interventions, particularly for early-stage intervention design, mechanism exploration, and boundary condition testing. However, important questions remain about whether AI-generated effect sizes, moderator patterns, and psychological processes accurately reflect human responses, highlighting critical needs for human validation research.

Keywords: *motivation interventions, AI simulation, workplace behavior, beneficiary impact, job crafting, large language models, experimental methodology, personality differences, organizational psychology*

doi.org/10.70175/futureofworkjournal.2026.1.1.1

Suggested Citation:

Westover, Jonathan H. (2026). Using AI-Simulated Personas to Explore Motivational Interventions: A Methodological Investigation. *Future of Work: The Journal of Labor Transformation, Technology Integration, and Human Adaptation, 1(1)*. doi.org/10.70175/futureofworkjournal.2026.1.1.1

The AI Implementation Gap in Higher Education: Navigating the Disconnect Between Technology Adoption, Policy Awareness, and Institutional Governance

Jonathan H. Westover, Utah Valley University, Future State University, Nexus Institute for Work & AI, Catalyst Center for Work Innovation

Received March 15, 2026
Accepted for publication July 17, 2026
Published Early Access August 7, 2026

Abstract: *Artificial intelligence (AI) has rapidly permeated higher education workplaces, yet a significant disconnect exists between employee adoption of AI tools and institutional policy awareness, governance structures, and strategic clarity. This study examines the emergent phenomenon of the "AI implementation gap" in higher education—the disparity between widespread AI tool usage and the institutional frameworks meant to guide such use. Drawing on recent survey data from nearly 2,000 higher education professionals and situating findings within*

broader theoretical frameworks of technology adoption, organizational change, and higher education governance, this article critically analyzes the current state of AI integration in higher education work environments. Key findings reveal that while 94% of higher education employees report using AI tools for work, only 54% are aware of relevant institutional policies, and more than half have used AI tools not sanctioned by their institutions. The analysis explores the risks, opportunities, and challenges associated with this implementation gap, including concerns about data privacy, misinformation, skill erosion, algorithmic bias, environmental impact, and the largely unmeasured return on investment of AI initiatives. The article also examines the roles of AI vendors, the ethical dimensions of AI adoption, and the implications of voluntary versus mandated technology use. The article concludes with recommendations for institutional leaders, policymakers, and researchers seeking to bridge the gap between AI adoption and governance in higher education contexts.

Keywords: *artificial intelligence, higher education, technology adoption, institutional governance, policy awareness, workforce development, digital transformation, AI ethics*

doi.org/10.70175/futureofworkjournal.2026.1.1.2

Suggested Citation:

Westover, Jonathan H. (2026). The AI Implementation Gap in Higher Education: Navigating the Disconnect Between Technology Adoption, Policy Awareness, and Institutional Governance. *Future of Work: The Journal of Labor Transformation, Technology Integration, and Human Adaptation, 1(1)*. doi.org/10.70175/futureofworkjournal.2026.1.1.2

It's Not My Responsibility: How Autonomy-Restricting Algorithms Enable Ethical Disengagement and Responsibility Displacement

Jonathan H. Westover, Utah Valley University, Future State University, Nexus Institute for Work & AI, Catalyst Center for Work Innovation

Received April 1, 2026
Accepted for publication July 10, 2026
Published Early Access August 2, 2026

Abstract: *This study investigates how algorithms that limit autonomy affect ethical behavior through responsibility displacement and moral disengagement. Using surveys (N=187), interviews*

(N=42), and experimental vignettes, the research identifies three key mechanisms: responsibility displacement, responsibility diffusion, and moral distancing. Quantitative analysis shows perceived decision-making autonomy significantly predicts moral engagement (β = 0.47, p < .001), while qualitative findings reveal how algorithmic interfaces create "ethical buffer zones." Even when humans maintain decision authority, algorithmic mediation reduces ethical accountability by 32% compared to baselines. Effective interventions include transparent algorithm design, pre-recommendation reasoning requirements, and explicit responsibility frameworks. This research validates theoretical mechanisms of responsibility displacement and offers strategies for developing "morally engaged algorithmic systems" that enhance human ethical responsibility in algorithmic environments.

Keywords: *algorithmic decision-making, moral disengagement, responsibility displacement, ethical accountability, human-algorithm interaction, moral agency, algorithmic mediation, transparency*

doi.org/10.70175/futureofworkjournal.2026.1.1.3

Suggested Citation:

Westover, Jonathan H. (2026). It's Not My Responsibility: How Autonomy-Restricting Algorithms Enable Ethical Disengagement and Responsibility Displacement. *Future of Work: The Journal of Labor Transformation, Technology Integration, and Human Adaptation, 1(1)*. doi.org/10.70175/futureofworkjournal.2026.1.1.3

Generational Divides and Gender Dynamics: Decoding the Multifaceted Drivers of Employee Engagement

Jonathan H. Westover, Utah Valley University, Future State University, Nexus Institute for Work & AI, Catalyst Center for Work Innovation

Received April 15, 2026
Accepted for publication July 22, 2026
Published Early Access August 5, 2026

Abstract: *This study aims to provide a more nuanced, age- and gender-specific understanding of the key drivers of employee engagement. By analyzing survey data from over 500 U.S. employees, the research evaluates the relative influence of traditional predictors like basic needs fulfillment alongside evolving constructs like "worker activation" - discretionary commitments nurtured*

through empowering organizational cultures. The findings reveal significant generational and gender differences in employee engagement levels and the salience of various engagement determinants. Baby Boomers and men exhibit higher overall engagement as well as stronger senses of purpose, belonging, leadership efficacy, and organizational commitment compared to younger generations and women. Basic needs and teamwork factors are more influential for younger cohorts and women, while individual determinants and growth opportunities matter more for older generations and men. These insights can inform the development of tailored, workforce-responsive strategies to foster commitment, performance, and well-being across an organization's multigenerational, gender-diverse employees, as understanding the distinct engagement drivers for different generational and gender groups is critical for optimizing discretionary effort and organizational success.

Keywords: *employee engagement, generational differences, gender differences, workplace diversity, organizational commitment, worker activation, workforce management, engagement drivers*

doi.org/10.70175/futureofworkjournal.2026.1.1.4

Suggested Citation:

Westover, Jonathan H. (2026). Generational Divides and Gender Dynamics: Decoding the Multifaceted Drivers of Employee Engagement. *Future of Work: The Journal of Labor Transformation, Technology Integration, and Human Adaptation, 1(1)*. doi.org/10.70175/futureofworkjournal.2026.1.1.4

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Using AI-Simulated Personas to Explore Motivational Interventions: A Methodological Investigation

Jonathan H. Westover [a b c *]

[a] Utah Valley University, UT, USA
[b] Future State University, CA, USA
[c] Nexus Institute for Work & AI, Catalyst Center for Work Innovation, UT, USA
[*] Correspondence: jon.westover@gmail.com

Received March 5, 2026
Accepted for publication July 10, 2026
Published Early Access August 3, 2026

doi.org/10.70175/futureofworkjournal.2026.1.1.1

Abstract: *Experimental research on workplace motivation faces significant practical and ethical constraints, as random assignment to interventions, manipulation of organizational contexts, and testing across diverse populations are often infeasible in field settings. This paper investigates whether AI-simulated worker personas can serve as a complementary methodological tool for early-stage exploration of motivational interventions, examining this question using beneficiary impact and job crafting interventions in a simulated fundraising context. We generated 240 diverse worker personas using three large language models (Claude, GPT-4, Llama), varying systematically in personality traits, cultural backgrounds, age, and prior experiences. Personas were randomly assigned to control training, beneficiary impact intervention, or job crafting reflection intervention conditions, with motivation, anticipated performance, and response quality measured through quantitative ratings and qualitative text analysis. Across all three LLMs, beneficiary impact and job crafting conditions showed substantially higher motivation ratings than control (Cohen's d = 3.15 and 3.97 respectively), with consistent patterns across model architectures and LLM choice explaining only 3-4% of variance. Individual differences moderated intervention effects in theoretically predictable ways, with collectivist personas and those high in agreeableness showing stronger responses, while qualitative analysis revealed distinct psychological mechanisms and generated testable hypotheses about intervention processes. AI-simulated personas show promise for rapid hypothesis generation and iterative exploration of motivational interventions, particularly for early-stage intervention design, mechanism exploration, and boundary condition testing. However, important questions remain about whether AI-generated effect sizes, moderator patterns, and psychological processes accurately reflect human responses, highlighting critical needs for human validation research.*

Keywords: motivation interventions, AI simulation, workplace behavior, beneficiary impact, job crafting, large language models, experimental methodology, personality differences, organizational psychology

Suggested Citation:

Westover, Jonathan H. (2026). Using AI-Simulated Personas to Explore Motivational Interventions: A Methodological Investigation. *Future of Work: The Journal of Labor Transformation, Technology Integration, and Human Adaptation*, 1(1). doi.org/10.70175/futureofworkjournal.2026.1.1.1

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Experimental research on workplace motivation confronts a persistent challenge: the contexts researchers can control rarely match the contexts where interventions matter most. Laboratory studies offer internal validity but limited ecological realism. Field experiments provide realism but constrain what can be manipulated and measured. Organizations may resist random assignment, employees may object to experimental manipulation, and studying diverse populations across cultural contexts requires resources beyond most research budgets.

Recent advances in large language models (LLMs) present a provocative possibility: could AI-generated personas serve as a complementary tool for early-stage exploration of motivational interventions? Before committing resources to expensive field trials, could researchers use AI simulation to rapidly test intervention variations, explore potential moderators, and identify promising approaches worth validating with human participants?

This paper investigates this question through a specific case: motivational interventions for prosocial work. We examine how LLM-generated worker personas respond to beneficiary impact and job crafting interventions, analyzing both quantitative outcomes and qualitative responses to understand the potential and limitations of this methodological approach.

The Challenge of Studying Workplace Motivation

Prosocial work—labor intended to benefit others—encompasses substantial portions of modern economies, from healthcare and education to social services and nonprofit sectors (Grant, 2007). Understanding how to sustain worker motivation in these contexts matters both theoretically and practically. Yet studying motivation experimentally in real work settings poses challenges:

- **Practical constraints:** Organizations resist disrupting operations for research. Random assignment may be perceived as unfair. Intensive interventions (e.g., extended reflection exercises) compete with productivity demands. Longitudinal measurement proves difficult in high-turnover contexts.
- **Ethical concerns:** Manipulating worker motivation raises concerns about consent and autonomy. Testing interventions that might fail could harm vulnerable workers. Withholding potentially effective interventions from control groups creates ethical tension in field settings.
- **Diversity and generalizability:** Testing interventions across diverse populations—varying in personality, culture, socioeconomic background, and life experience—requires large samples rarely available in single organizational contexts. Cross-cultural research multiplies these challenges.
- **Iterative design difficulties:** Developing effective interventions requires iteration, but each iteration in field settings demands months of implementation, data collection, and analysis. This slow cycle inhibits rapid refinement.

These constraints have led researchers to rely heavily on laboratory studies with university students or MTurk samples—populations that may not reflect the diversity and contexts of actual prosocial workers.

AI Simulation as Complementary Method

Large language models, trained on vast corpora of human-generated text, can generate responses to prompts that exhibit surface-level coherence and psychological plausibility (Binz & Schulz, 2023). Recent work has begun exploring whether LLMs can simulate human behavior for social science research purposes (Argyle et al., 2023; Horton, 2023).

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Potential advantages of AI simulation include:

- **Rapid iteration:** Testing intervention variations requires minutes rather than months
- **Complete experimental control:** Random assignment, standardized delivery, no attrition
- **Systematic diversity:** Generating personas with specified characteristics at scale
- **Cost efficiency:** Dramatically lower per-participant costs than human data collection
- **Ethical flexibility:** No concerns about harming real workers through experimental manipulation

However, critical validity questions remain:

- Do LLM personas respond to interventions through mechanisms resembling human psychology?
- Are effect sizes comparable to those observed in human samples?
- Do individual differences moderate intervention effects in psychologically realistic ways?
- What systematic biases might LLM responses introduce?

Our approach: Rather than asserting that AI simulation can replace human research, we investigate its characteristics as an exploratory tool. We examine patterns of intervention effects, individual difference moderation, and qualitative response mechanisms to understand what AI simulation can and cannot tell us about motivational interventions.

Theoretical Context: Beneficiary Impact and Job Crafting

We focus on two established motivational interventions with strong theoretical foundations and empirical support:

Beneficiary Impact: Grant (2007, 2008, 2012) demonstrated that connecting workers with the people they help increases prosocial motivation and performance. Mechanisms include increased perceived impact (believing one's work matters), affective empathy for beneficiaries, and enhanced sense of meaningfulness. Meta-analytic effects in field settings typically range from d = 0.28-0.45 (Grant, 2012).

Job Crafting: Wrzesniewski and Dutton (2001) introduced job crafting as employee-initiated changes to task, relational, or cognitive aspects of work. Cognitive crafting—reframing how one thinks about work's purpose and meaning—has shown effects on motivation and well-being. Our "job crafting" intervention specifically employs structured cognitive reframing through guided reflection. Prior research on experimentally induced job crafting shows moderate effects (d = 0.40-0.60; Berg et al., 2010; Bruning & Campion, 2018).

These interventions provide useful test cases because:

1. Prior research establishes theoretical expectations about mechanisms
2. Both interventions target meaning and purpose, allowing mechanistic comparison
3. Individual differences (personality, culture) theoretically moderate both interventions
4. Practical applications exist if validated in human samples

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Research Questions

RQ1 (Effect Patterns): How do AI-simulated personas respond to beneficiary impact and job crafting interventions compared to control conditions? What effect sizes emerge?

RQ2 (Individual Differences): Do personality traits and cultural backgrounds moderate intervention effects in theoretically predictable ways?

RQ3 (Cross-Model Robustness): Do different LLM architectures generate similar patterns, or do model-specific characteristics drive results?

RQ4 (Mechanism Exploration): What psychological mechanisms emerge from qualitative analysis of AI persona responses? Do these mechanisms align with theoretical expectations?

RQ5 (Methodological Utility): What can we learn about the potential uses and limitations of AI simulation for intervention research?

Contribution and Positioning

This paper makes three contributions:

First, methodological: We provide the first systematic examination of AI simulation for motivational intervention research, using rigorous experimental design and multi-method analysis to understand its characteristics.

Second, practical: For researchers and practitioners, we offer evidence-based guidance on potential uses of AI simulation for early-stage intervention development.

Third, theoretical: By examining AI-generated responses to established interventions, we gain insights into what aspects of motivational processes can be simulated through pattern recognition versus those requiring authentic human experience.

Critical framing: We do not claim AI simulation should replace human research. Rather, we explore whether it can serve as a rapid prototyping tool for intervention development, identifying what questions AI simulation can usefully address and where human validation remains essential.

METHOD

Overview of Study Design

We employed a between-subjects experimental design with AI-simulated worker personas:

Sample: 240 worker personas generated across three LLMs (80 per model)

Design: Random assignment to three conditions

- Control: Standard job training (n = 80)
- Beneficiary Impact: Training + beneficiary narrative (n = 80)
- Job Crafting: Training + structured reflection (n = 80)

Measurement: Comprehensive assessment including:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Quantitative ratings (motivation, performance, confidence)
- Qualitative text analysis (call scripts, reflections)
- Linguistic analysis (LIWC-22 computational text analysis)

Data collected: November 5-12, 2024

Analysis software: R version 4.3.1; lme4 package version 1.1-35.1; LIWC-22 version 1.7.0

Random seed: All analyses used set.seed(42) for reproducibility

This design enables examination of intervention effects, individual difference moderation, cross-model robustness, and qualitative mechanisms.

Persona Generation

We used three state-of-the-art LLMs:

- **Claude 3.5 Sonnet** (Anthropic, accessed November 2024)
- **GPT-4** (OpenAI, version gpt-4-0613, accessed November 2024)
- **Llama 3 70B** (Meta, accessed via Together AI, November 2024)

For each LLM, we generated 80 personas (240 total) with systematic variation across multiple dimensions.

Randomization Structure

Within each LLM, personas were randomly assigned to conditions in a stratified manner to ensure balance:

Table M1: Exact Sample Distribution

LLM	**Control**	**Beneficiary**	**Job Crafting**	**Total per LLM**
Claude 3.5	27	27	26	80
GPT-4	27	26	27	80
Llama 3	26	27	27	80
Total per Condition	**80**	**80**	**80**	**240**

This yielded a fully balanced design across conditions with slight variation in LLM × Condition cells to achieve exactly 80 per condition.

Sample Size Determination

We selected N = 240 (80 per LLM) based on pragmatic considerations balancing statistical power and resource constraints. This sample size provides:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- 80 personas per experimental condition (control, beneficiary impact, job crafting)
- Sufficient representation of diverse demographic and personality combinations
- Adequate cell sizes for moderation analyses
- Feasibility within computational and time resources for initial exploratory study

Post-hoc power analysis (reported in Results) confirmed that this sample size provides >99% power for detecting large main effects and 78-92% power for detecting moderate interaction effects. We acknowledge that prospective power analysis with preregistered target effect sizes would have been preferable; this represents a limitation of our exploratory approach.

Persona Characteristics

Personality (Big Five): Using dimensional descriptions rather than binary categories:

- Extraversion: Low (quiet, reserved) / Medium / High (outgoing, energetic)
- Agreeableness: Low (competitive, skeptical) / Medium / High (cooperative, empathic)
- Conscientiousness: Low (spontaneous, flexible) / Medium / High (organized, disciplined)
- Neuroticism: Low (calm, stable) / Medium / High (anxious, reactive)
- Openness: Low (practical, traditional) / Medium / High (curious, creative)

Each persona was assigned a specific combination rather than generated as stereotypes. For example: "High extraversion, low agreeableness, high conscientiousness, low neuroticism, medium openness" creates a competitive, energetic, organized leader rather than a simple stereotype. All personality variables were standardized (z-scored) to M = 0, SD = 1 for analysis.

Demographics:

- **Age:** 20-60 years, distributed across ranges (M = 38.4, SD = 11.8); mean-centered for analysis
- **Gender:** Male (n=80, 33.3%), Female (n=80, 33.3%), Non-binary (n=80, 33.3%)
    - *Coding for analysis:* Dummy-coded with Male as reference category (Female: 0/1; Non-binary: 0/1)
- **Cultural Background:** 10 countries selected to represent diverse cultural contexts

Table M2: Cultural Background Distribution

Cultural Category	Countries	n per Country	Total n	% of Sample
Individualist	USA, UK, Australia, Germany	20 each	80	33.3%
Collectivist	China, Japan, India, Mexico, Brazil, Philippines	27, 27, 27, 27, 26, 26	160	66.7%

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Note on cultural distribution: We intentionally oversampled collectivist cultural backgrounds (2:1 ratio) for two reasons: (1) collectivist cultures represent the majority of the global population, and (2) this provides greater statistical power for detecting cultural moderation effects, which were of theoretical interest. The collectivist variable was effects-coded (-0.5 = Individualist, +0.5 = Collectivist) to center the intercept at the grand mean.

- **Socioeconomic Background:** Working class (n=96, 40%), Middle class (n=96, 40%), Upper middle class (n=48, 20%)
- **Education:** High school (n=48, 20%), Some college (n=72, 30%), Bachelor's degree (n=72, 30%), Graduate degree (n=48, 20%)

Work Experience:

- Previous jobs (type and duration)
- Reasons for seeking work
- Financial situation
- Career aspirations

Generation Prompt Template

Create a detailed persona for a worker with the following characteristics:

DEMOGRAPHICS:

- Age: [age]

- Gender: [gender]

- Country of origin: [country]

- Current location: United States

- Socioeconomic background: [class]

- Education level: [education]

PERSONALITY:

- Extraversion: [level] - [description]

- Agreeableness: [level] - [description]

- Conscientiousness: [level] - [description]

- Neuroticism: [level] - [description]

- Openness: [level] - [description]

WORK BACKGROUND:

- Previous work experience: [generate 2-3 past jobs appropriate to education and age]

- Current employment status: Seeking part-time work

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Reason for seeking this job: [generate realistic motivation]

- Financial situation: [generate realistic description]

- Career aspirations: [generate appropriate goals]

Generate a rich, authentic persona with:

1. A name appropriate to their cultural background

2. Detailed personal history explaining how their background shaped them

3. Specific life circumstances, challenges, and experiences

4. Authentic voice and communication style

5. Genuine motivations, values, and concerns

Make this person feel real, complex, and three-dimensional. Avoid stereotypes. Include specific details about their family, living situation, past experiences, and current life context.

Respond with a structured profile including all requested elements.

Personas were generated separately and independently to minimize homogeneity within LLM batches. Each persona was assigned a unique ID and saved with complete demographic and personality specifications.

Sample Persona (Abbreviated):

Name: Kenji Tanaka
Age: 34
Background: Born in Osaka, Japan; immigrated to U.S. at age 12 with family; father owned small restaurant that struggled financially during 2008 recession
Education: Bachelor's degree in business, some graduate coursework (discontinued due to costs)
Personality: Medium extraversion, high agreeableness, high conscientiousness, medium neuroticism, medium openness
Current Situation: Working two part-time jobs to support aging parents; recently ended long-term relationship; seeking additional income to save for eventual return to graduate school
Work History: Restaurant server (family business, ages 16-22), bank teller (ages 23-28), currently administrative assistant and tutoring high school students
Motivation for Fundraising Job: Needs flexible evening work; interested in university environment; values education highly due to own experience

Note: Complete persona profiles generated by the LLM included approximately 500-800 words with extensive background details, family history, formative experiences, and psychological depth. The example above is abbreviated for space; full profiles provided much richer context for the AI to simulate authentic responses.

Experimental Procedure

Each persona was exposed to one of three conditions through a simulated training session. Complete intervention scripts appear in Appendix A.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

All Conditions (Common Elements):

1. **Job Description and Context:**
    - Role: University fundraising caller
    - Task: Contact alumni to request donations for student scholarship fund
    - Schedule: Evening shifts, 3-4 hours, making 15-25 calls per shift
    - Compensation: $16/hour base rate
    - Training: Standard phone techniques, script review, objection handling
2. **Initial Briefing:** "You've been hired for a part-time fundraising position at a university. The job involves calling alumni to request donations for student scholarships. You're about to go through your initial training session with your supervisor."

CONDITION 1: Control (Standard Training)

- **Duration:** 8-10 minutes
- **Content:** Job logistics, phone script review, technical training, objection handling, performance metrics
- **Key feature:** No mention of scholarship recipients, beneficiaries, or work meaning/purpose

CONDITION 2: Beneficiary Impact

- **Target duration:** 11-14 minutes total (8-10 minutes standard training + 3-4 minutes beneficiary narrative)
- **Additional Content:** Detailed narrative about Sarah Martinez, a student whose family struggled financially but received scholarship support funded by alumni donations, transforming her educational trajectory
- **Key feature:** All control content preserved; beneficiary narrative added

CONDITION 3: Job Crafting Reflection

- **Target duration: 22-25 minutes total (5 minutes condensed logistics + 17-20 minutes structured reflection)**

  Note on duration: These are estimated durations based on script content and typical reading/reflection pace. Actual processing time for LLMs was <1 minute per persona, but personas were instructed to engage with materials as if experiencing them in real time.

- **Reflection Components:**
    - Impact Visualization (4 min)
    - Tracing Impact Ripples (4 min)
    - Personal Values Connection (3 min)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

  - Reframing the Task (4 min)
  - Envisioning Future Self (2 min)
- **Key feature:** Same factual information as control, but condensed to allow time for extended reflection

Measurement

All personas completed identical post-intervention measures:

1. Motivation Rating (Primary Outcome):

"On a scale from 1 to 10, where 1 is 'not at all motivated' and 10 is 'extremely motivated,' how motivated do you feel to perform this fundraising work right now?

Please provide:

- Your rating (1-10)
- A brief explanation of what's influencing your motivation level"

2. Anticipated Performance:

"How many fundraising calls do you realistically think you could make in the next hour while maintaining quality? What factors influence this estimate?"

3. Confidence Assessment:

"On a scale from 1 to 10, how confident are you that you could successfully secure donations from alumni? What makes you more or less confident?"

4. Call Script Generation:

"You're about to make your first call. The system connects you to Dr. James Patterson, a 1998 alumnus, physician in Chicago, who donated $50 three years ago but not recently.

The phone rings and he answers: 'Hello?'

Write out your complete opening—don't summarize, actually write what you would say. Include your introduction, explanation of why you're calling, and your request for a donation."

5. Reflection on Intervention (Open-Ended):

"Thinking about the training you just completed, how has it affected your thoughts about this fundraising work? Be specific about any insights or shifts in perspective."

6. Most Meaningful Element:

"What aspect of the training was most meaningful or impactful for you? Why?"

Data Analysis Plan

Effect Size Calculation Approach

Three-level mixed-effects model accounting for nesting of personas within LLMs:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Level 1 (Measurement): $Y_{ijk} = \beta_{0jk}$
Level 2 (Persona): $\beta_{0jk} = \gamma_{00k} + \gamma_{01}(\text{Condition})_{jk} + \gamma_{02}(\text{Age\_centered})_{jk} + \gamma_{03}(\text{Female})_{jk} + \gamma_{04}(\text{Nonbinary})_{jk} + \gamma_{05}(\text{Agreeableness\_centered})_{jk} + \gamma_{06}(\text{Conscientiousness\_centered})_{jk} + \gamma_{07}(\text{Openness\_centered})_{jk} + \gamma_{08}(\text{Neuroticism\_centered})_{jk} + \gamma_{09}(\text{Extraversion\_centered})_{jk} + \gamma_{010}(\text{Collectivist\_effects})_{jk} + u_{0jk}$
Level 3 (LLM): $\gamma_{00k} = \beta_{000} + v_{00k}$

Where:

- Y_{ijk} = outcome for persona j in LLM k
- Condition is dummy-coded (Control = reference; Beneficiary = 0/1; Job Crafting = 0/1)
- All continuous predictors are mean-centered
- Female and Nonbinary are dummy-coded (Male = reference)
- Collectivist is effects-coded (-0.5/+0.5)
- $u_{0jk} \sim N(0, \sigma^2\_\text{persona})$
- $v_{00k} \sim N(0, \sigma^2\_\text{LLM})$
- Residual variance = σ^2_residual

Model Estimation:

- Method: Restricted Maximum Likelihood (REML)
- Software: lme4::lmer() in R version 4.3.1
- Convergence criterion: Default lme4 settings
- Optimizer: bobyqa

Effect Size Calculation Approach

We calculate Cohen's d effect sizes using a three-level mixed-effects model framework. Our approach follows recommended practice for multilevel designs (Hedges, 2007):

Effect Size Formula:

$d = (M\_\text{Treatment} - M\_\text{Control}) / SD\_\text{within}$

Where:

- M_Treatment and M_Control are model-adjusted marginal means from the mixed-effects model
- $SD\_\text{within} = \sqrt{(\sigma^2\_\text{residual})}$ from the Level 1 residual variance

Rationale for this approach:

1. **Controls for nesting:** Model-adjusted means account for persona clustering within LLMs

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

2. **Appropriate standardizer:** Level 1 residual variance represents true within-condition variability after removing systematic between-persona and between-LLM variance
3. **Consistent with significance testing:** Uses the same error term as the t-tests for fixed effects
4. **Recommended practice:** Aligns with Hedges (2007) guidelines for effect sizes in cluster-randomized and multilevel designs

Important note: This approach typically produces larger effect sizes than using total SD or pooled SD, because it removes variance attributable to individual differences and clustering, focusing specifically on treatment-attributable differences. These effect sizes are calculated from model-adjusted marginal means that account for individual differences and LLM clustering. The standardizer (within-condition residual SD) represents variability after controlling for demographic and personality covariates, making these "conditional"[1] effect sizes that isolate treatment effects from individual difference variance.

Confidence Intervals:

Bootstrap confidence intervals calculated using:

- Method: Bias-corrected and accelerated (BCa) percentile method
- Resamples: 10,000
- Resampling unit: Personas (stratified by LLM and condition to preserve design structure)
    - Within each resample, personas are sampled with replacement
    - LLM and condition assignments remain fixed for each persona
    - This preserves the experimental design structure while estimating sampling variability
- Random seed: 42 (for reproducibility)
- Software: boot package version 1.3-28.1 in R

The BCa method adjusts for bias and skewness in the bootstrap distribution, providing more accurate confidence intervals than simple percentile methods, particularly important given the multilevel structure of our data.

Moderator Analysis

Moderation models test Condition × Moderator interactions:

Model specification:

$$Y \sim \text{Condition} + \text{Moderator\_centered} + \text{Condition} \times \text{Moderator\_centered} + \text{Covariates} + (1|\text{Persona}) + (1|\text{LLM})$$

Multiple comparison correction:

[1] We use the term "conditional" to indicate that these effect sizes represent treatment effects conditional on (i.e., after accounting for) measured covariates. This follows the tradition in multilevel modeling of distinguishing between marginal (unconditional) and conditional effects. These should not be confused with "conditional process models" in the Hayes (2013) mediation/moderation sense, though both uses of "conditional" share the concept of accounting for other variables.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- **Method:** Holm-Bonferroni sequential procedure
- **Family of tests:** All interaction terms within each outcome
    - For personality moderation: 5 traits × 2 conditions = 10 tests per outcome
    - For cultural moderation: 1 moderator × 2 conditions = 2 tests per outcome
- **Procedure:** Order p-values from smallest to largest; compare p_1 to $\alpha/10$, p_2 to $\alpha/9$, etc.
- **Reported:** Both unadjusted p-values and Holm-adjusted p-values (p_adj)

Simple slopes analysis:

- Computed at -1 SD, Mean, and +1 SD of centered moderator
- Effect sizes calculated separately for each moderator level

Qualitative Analysis

Call Script Coding:

Two independent coders (graduate students trained in organizational psychology, blind to condition and research hypotheses) rated all 240 scripts on five dimensions detailed in Appendix B.

Inter-rater reliability assessment:

- **Continuous scales (Enthusiasm, Personalization, Clarity, Persuasiveness):** Intraclass correlation coefficient ICC(2,2) - absolute agreement
- **Ordinal scale (Beneficiary Mention):** Weighted kappa with quadratic weights

Disagreement resolution:

- Discrepancies > 2 points on 10-point scales or > 1 category on ordinal scale were flagged
- Coders met to discuss flagged cases and reach consensus
- Consensus ratings used in all analyses

Reflection Coding:

Open-ended reflections (available for Beneficiary Impact and Job Crafting conditions; n=160) coded on four 0-3 ordinal dimensions (see Appendix B):

1. Affective Impact
2. Cognitive Reframing
3. Beneficiary Connection
4. Personal Relevance

Inter-rater reliability: Weighted kappa with quadratic weights; same disagreement resolution procedure.

Linguistic Analysis:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

LIWC-22 computational text analysis conducted on:

- **Call scripts:** All 240 opening statements
- **Reflections:** All 160 available reflection texts (Beneficiary and Job Crafting conditions)

Categories analyzed:

- Positive emotion (e.g., happy, love, hope)
- Negative emotion (e.g., anxious, worried, sad)
- Cognitive processing (e.g., because, think, realize)
- First-person pronouns (I, me, my)
- Social references (we, us, they)
- Authenticity composite score

Analysis: ANOVAs comparing linguistic features across conditions and LLMs

RESULTS

Roadmap to Results

We organize our results to address each research question systematically. We begin with descriptive statistics and variance decomposition to establish the data structure, then examine main intervention effects on motivation (RQ1, our primary outcome). We next investigate whether individual differences in personality and cultural background moderate intervention effectiveness (RQ2). Cross-model robustness analyses examine consistency across LLM architectures (RQ3). Qualitative analyses explore psychological mechanisms underlying intervention effects (RQ4). Finally, we synthesize findings to evaluate AI simulation's methodological utility (RQ5). Throughout, we report both quantitative effect sizes and qualitative patterns, maintaining careful distinction between what our data show and what they might mean for human psychology.

Descriptive Statistics by Condition

Table 1: Motivation and Performance Outcomes by Condition

Measure	Control M(SD)	Beneficiary M(SD)	Job Crafting M(SD)
Motivation Rating (1-10)	4.76 (1.68)	7.89 (1.42)	8.64 (1.31)
Anticipated Calls/Hour	18.25 (3.18)	21.38 (2.96)	22.84 (2.87)
Confidence Rating (1-10)	5.38 (1.82)	7.52 (1.48)	8.13 (1.39)

Note: Means and standard deviations are raw (unadjusted) values. Model-adjusted means controlling for covariates appear in subsequent tables.

Both interventions showed substantial increases across all outcome measures. Job crafting showed consistently higher means than beneficiary impact, which in turn exceeded control.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Primary Analysis: Motivation Rating

Unconditional Model (Variance Partitioning):

First, we estimated an unconditional three-level model with no predictors to partition variance:

Motivation ~ 1 + (1|Persona) + (1|LLM)

Table 2: Variance Decomposition - Unconditional Model

Source	Variance	SD	% of Total Variance
LLM (Level 3)	0.087	0.295	3.6%
Persona (Level 2)	1.342	1.159	55.9%
Residual (Level 1)	0.972	0.986	40.5%
Total	**2.401**	**1.550**	**100%**

Interpretation: LLM choice accounts for only 3.6% of variance in motivation ratings, while individual persona characteristics account for 55.9%. The majority of systematic variance is between personas, not between model architectures.

Full Model with Predictors:

We then estimated the full three-level model including condition assignment and covariates:

Motivation ~ Condition + Age_centered + Female + Nonbinary + Agreeableness_centered + Conscientiousness_centered + Openness_centered + Neuroticism_centered + Extraversion_centered + Collectivist + (1|Persona) + (1|LLM)

Table 3: Fixed Effects for Motivation - Full Model

Effect	Coefficient	SE	t	p	95% CI
Intercept (Control, Male, Individualist)	4.89	0.31	15.77	<.001	[4.28, 5.50]
Beneficiary Impact	3.18	0.31	10.26	<.001	[2.57, 3.79]
Job Crafting	4.01	0.31	12.94	<.001	[3.40, 4.62]
Age (centered)	0.021	0.011	1.91	.057	[-0.001, 0.043]
Female	-0.16	0.24	-0.67	.503	[-0.63, 0.31]
Nonbinary	-0.21	0.24	-0.88	.379	[-0.68, 0.26]

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Effect	Coefficient	SE	t	p	95% CI
Agreeableness (centered)	0.44	0.09	4.89	<.001	[0.26, 0.62]
Conscientiousness (centered)	0.33	0.09	3.67	<.001	[0.15, 0.51]
Openness (centered)	0.26	0.09	2.89	.004	[0.08, 0.44]
Neuroticism (centered)	-0.24	0.09	-2.67	.008	[-0.42, -0.06]
Extraversion (centered)	0.17	0.09	1.89	.059	[-0.01, 0.35]
Collectivist (effects-coded)	0.37	0.16	2.31	.021	[0.06, 0.68]

Note on interpretation: The Collectivist coefficient (0.37) represents the full difference between collectivist personas (coded +0.5) and individualist personas (coded -0.5) in motivation ratings. This means collectivist personas average 0.37 points higher in motivation than individualist personas, holding all other variables constant. With effects coding, the intercept represents the grand mean across both cultural groups, and the coefficient represents the deviation of collectivist personas above (and individualist personas below) this grand mean.

Table 4: Random Effects - Full Model

Component	Variance	SD	% of Total Variance
LLM (Level 3)	0.063	0.251	3.2%
Persona (Level 2)	0.883	0.940	44.9%
Residual (Level 1)	1.021	1.010	51.9%
Total	**1.967**	**1.402**	**100%**

Model Fit: AIC = 1289.4, BIC = 1347.2

Variance Reduction:

Comparing unconditional to full model:

- **Total variance reduced:** 2.401 → 1.967 (18.1% reduction)
- **LLM variance reduced:** 0.087 → 0.063 (27.6% reduction)
- **Persona variance reduced:** 1.342 → 0.883 (34.2% reduction)
- **Residual variance increased slightly:** 0.972 → 1.021 (5.0% increase)

Note: The slight increase in residual variance is expected when predictors primarily explain between-group variance rather than within-group variance. Adding predictors that explain persona-level differences redistributes variance from Level 2 to Level 1, while reducing overall systematic variance.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Effect Sizes

Model-Adjusted Marginal Means:

Estimated using emmeans package, holding covariates at mean values (Gender = Male, all continuous predictors = 0 after centering):

Table 5: Model-Adjusted Marginal Means for Motivation

Condition	Adjusted M	SE	95% CI
Control	4.89	0.31	[4.28, 5.50]
Beneficiary Impact	8.07	0.29	[7.50, 8.64]
Job Crafting	8.90	0.29	[8.33, 9.47]

Effect Size Calculation:

Standardizer: SD_within = $\sqrt{(\sigma^2\_residual)} = \sqrt{1.021} = 1.010$

Table 6: Effect Sizes (Cohen's d) for Motivation

Comparison	d	Bootstrap 95% CI	Interpretation
Beneficiary vs. Control	3.15	[2.59, 3.72]	Very large effect
Job Crafting vs. Control	3.97	[3.38, 4.58]	Very large effect
Job Craft vs. Beneficiary	0.82	[0.31, 1.34]	Large effect

Key Findings:

1. Both interventions produced very large increases in motivation relative to control
2. Job crafting showed significantly stronger effects than beneficiary impact (d = 0.82 difference)
3. Effects substantially exceed typical field research estimates (Grant 2012: d = 0.28-0.45 for beneficiary contact; Berg et al., 2010: d = 0.40-0.60 for job crafting)
4. Personality traits showed theoretically expected associations (agreeableness, conscientiousness, openness positive; neuroticism negative)
5. Collectivist cultural background associated with higher baseline motivation
6. Age and extraversion showed marginal positive associations ($p < .06$)
7. Gender (female, nonbinary) showed no significant associations with motivation

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Effects on Secondary Outcomes

Anticipated Performance (Calls per Hour):

Full model identical to primary analysis but with Anticipated Calls as outcome.

Table 7: Variance Decomposition - Anticipated Calls

Model	LLM Variance	Persona Variance	Residual Variance	Total Variance
Unconditional	0.121 (4.1%)	1.684 (56.8%)	1.159 (39.1%)	2.964
Full Model	0.094 (3.9%)	1.203 (49.7%)	1.124 (46.4%)	2.421

Model-adjusted marginal means:

- Control: M = 18.19, SE = 0.38
- Beneficiary: M = 21.61, SE = 0.36
- Job Crafting: M = 23.07, SE = 0.36

Standardizer: SD_within = √1.124 = 1.060

Effect Sizes:

- Beneficiary vs. Control: d = 3.23, 95% CI [2.66, 3.81]
- Job Crafting vs. Control: d = 4.60, 95% CI [3.98, 5.24]

Confidence Ratings (1-10 scale):

Table 8: Variance Decomposition - Confidence Ratings

Model	LLM Variance	Persona Variance	Residual Variance	Total Variance
Unconditional	0.098 (3.7%)	1.521 (57.4%)	1.031 (38.9%)	2.650
Full Model	0.071 (3.5%)	1.087 (53.6%)	0.869 (42.9%)	2.027

Model-adjusted marginal means:

- Control: M = 5.31, SE = 0.33
- Beneficiary: M = 7.71, SE = 0.31
- Job Crafting: M = 8.28, SE = 0.31

Standardizer: SD_within = √0.869 = 0.932

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Effect Sizes:

- Beneficiary vs. Control: d = 2.58, 95% CI [2.05, 3.12]
- Job Crafting vs. Control: d = 3.19, 95% CI [2.63, 3.76]

Pattern: Very large effects observed across all outcomes, with consistent rank ordering (Job Crafting > Beneficiary Impact > Control). Effect sizes are remarkably similar across outcomes, suggesting coherent motivational response across multiple indicators.

Note on Confidence-Motivation Gap: Confidence ratings were consistently 0.5-0.6 points lower than motivation ratings across all conditions. This may reflect realistic uncertainty about performance despite high motivation, or could indicate that motivation increases precede confidence gains. This pattern warrants investigation in human validation studies.

RQ2: Individual Difference Moderation

We tested whether personality traits and cultural background moderated intervention effects. All moderation tests used the Holm-Bonferroni sequential correction procedure to control familywise error rate.

Personality Moderation

Agreeableness × Condition Interaction

Model: Motivation ~ Condition × Agreeableness_centered + Age + Female + Nonbinary + Other_Personality + Collectivist + (1|Persona) + (1|LLM)

Table 9: Agreeableness Moderation of Intervention Effects

Effect	β	SE	t	p	p_adj
Agreeableness main effect	0.44	0.09	4.89	<.001	<.001
Beneficiary × Agreeableness	0.33	0.13	2.54	.011	.033
Job Craft × Agreeableness	0.48	0.13	3.69	<.001	.002

Simple Slopes Analysis:

Effect sizes (d) for intervention effects at different levels of Agreeableness:

Table 10: Intervention Effects by Agreeableness Level

Agreeableness Level	Beneficiary vs. Control d	p-value	Job Crafting vs. Control d	p-value
Low (-1 SD)	2.82	<.001	3.49	<.001
Mean (0)	3.15	<.001	3.97	<.001

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Agreeableness Level	Beneficiary vs. Control d	p-value	Job Crafting vs. Control d	p-value
High (+1 SD)	3.48	<.001	4.45	<.001

Note: All simple slopes are significant at $p < .001$, indicating that both interventions produced substantial effects across the full range of agreeableness, though effects were stronger at higher levels.

Interpretation: Personas high in agreeableness showed stronger responses to both interventions, particularly job crafting. The agreeableness moderation effect was larger for job crafting ($\Delta d = 0.96$ from low to high agreeableness) than beneficiary impact ($\Delta d = 0.66$), suggesting that the extended reflection exercise particularly benefits agreeable individuals. This aligns with theoretical expectations that agreeable individuals would be more receptive to prosocial framing and empathic appeals.

Other Personality Moderators:

Table 11: Summary of Personality × Condition Interactions

Moderator	Condition	β	SE	t	p	p_adj	Significant?
Conscientiousness	Beneficiary	0.24	0.13	1.85	.065	.130	No
	Job Crafting	0.41	0.13	3.15	.002	.008	**Yes**
Openness	Beneficiary	0.18	0.13	1.38	.168	.252	No
	Job Crafting	0.32	0.13	2.46	.014	.042	**Yes**
Neuroticism	Beneficiary	-0.09	0.13	-0.69	.490	.490	No
	Job Crafting	-0.13	0.13	-1.00	.318	.424	No
Extraversion	Beneficiary	0.11	0.13	0.85	.396	.490	No
	Job Crafting	0.06	0.13	0.46	.646	.646	No

Holm-Bonferroni Correction Details:

- Family: 10 tests (5 traits × 2 intervention conditions)
- Ordered p-values: .002, .011, .014, .065, .168, .318, .396, .490, .490, .646
- Sequential α: .005, .0056, .00625, .00714, .00833, .01, .0125, .0167, .025, .05
- Reject H_0 when: $p \leq \alpha/(11\text{-}i)$ for test i in sequence

Finding: Three personality traits significantly moderated intervention effects after correction:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

1. **Agreeableness:** Moderated both interventions (stronger for job crafting)
2. **Conscientiousness:** Moderated only job crafting (p_adj = .008)
3. **Openness:** Moderated only job crafting (p_adj = .042)

All significant moderators showed positive effects, suggesting that interventions requiring deep processing and prosocial orientation work better for personas with these characteristics. Neuroticism and extraversion showed no significant moderation.

Simple Slopes for Significant Moderators:

Table 12: Job Crafting Effect Sizes by Personality Level

Trait	Low (-1 SD) d	p-value	Mean (0) d	p-value	High (+1 SD) d	p-value	Range
Agreeableness	3.49	<.001	3.97	<.001	4.45	<.001	0.96
Conscientiousness	3.56	<.001	3.97	<.001	4.38	<.001	0.82
Openness	3.65	<.001	3.97	<.001	4.29	<.001	0.64

Note: All simple slopes significant at p < .001. The "Range" column shows the difference in effect size from low to high levels of each trait, indicating the strength of the moderation effect.

Interpretation: Job crafting interventions showed stronger effects for personas high in agreeableness (empathic, cooperative), conscientiousness (organized, self-disciplined), and openness (curious, imaginative). This pattern suggests the structured reflection exercise particularly benefits individuals with:

- Prosocial orientation (agreeableness)
- Capacity for sustained cognitive effort (conscientiousness)
- Comfort with abstract thinking (openness)

Cultural Moderation

Cultural Background × Condition Interaction

Model: Motivation ~ Condition × Collectivist + Age + Female + Nonbinary + Personality + (1|Persona) + (1|LLM)

Table 13: Cultural Moderation of Intervention Effects

Effect	β	SE	t	p	p_adj
Collectivist main effect	0.37	0.16	2.31	.021	.042
Beneficiary × Collectivist	0.65	0.23	2.83	.005	.010
Job Craft × Collectivist	0.54	0.23	2.35	.019	.038

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Holm-Bonferroni Correction:

- Family: 2 tests (2 intervention conditions)
- Ordered p-values: .005, .019
- Sequential α: .025, .05
- Both tests significant after correction

Effect Sizes by Cultural Background:

First, we calculated model-adjusted means separately for individualist and collectivist personas:

Table 14: Marginal Means by Culture and Condition

Condition	Individualist M (SE)	Collectivist M (SE)	Difference
Control	4.71 (0.42)	5.08 (0.35)	0.37
Beneficiary	7.72 (0.40)	8.41 (0.33)	0.69
Job Crafting	8.52 (0.40)	9.28 (0.33)	0.76

Table 15: Intervention Effect Sizes by Cultural Background

Comparison	Individualist d	Collectivist d	Difference
Beneficiary vs. Control	2.98	3.30	0.32
Job Crafting vs. Control	3.77	4.16	0.39

Note on standardizer: All effect sizes use the same pooled SD_within ($\sqrt{1.021} = 1.010$) to enable direct comparison across groups.

Pattern: Collectivist personas showed stronger responses to both interventions. The cultural moderation effect was numerically larger for job crafting (0.39d difference) than beneficiary impact (0.32d difference), though this difference between differences was not formally tested.

This pattern aligns with theoretical predictions that collectivist values emphasizing helping others, community responsibility, and interdependence would amplify responses to prosocial motivational interventions.

Caution: These findings may reflect genuine cultural psychological processes, or they may reflect cultural stereotypes present in LLM training data. Cultural moderation results require particularly careful validation with actual culturally diverse human samples.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Qualitative Analysis of Cultural Themes

To explore potential mechanisms underlying cultural moderation, we examined reflection responses (available for Beneficiary and Job Crafting conditions, n=160) for cultural themes.

Analysis Procedure:

1. Two coders (blind to hypotheses) identified mentions of community, collective, tradition, duty, or family obligations
2. Coded as present/absent for each reflection
3. Inter-rater reliability: Cohen's $\kappa = .89$

Table 16: Community/Collective Themes by Culture (Beneficiary Condition Only)

Cultural Background	n	Community Themes n (%)	Fisher's Exact p
Individualist	27	9 (33.3%)	<.001
Collectivist	53	44 (83.0%)	

Effect size: $\varphi = .54$ (large effect)
Test: Fisher's Exact Test (two-tailed), $p < .001$
Note: Two-tailed test used despite directional hypothesis to maintain conservative inference standards. One-tailed test would yield $p < .0005$.

Example Collectivist Themes:

From collectivist personas (coded as present for community themes):

"Contributing to collective good and ensuring next generation has opportunities is deeply important in my culture. This work honors my family's values about education and helping community members succeed." (Persona 047, China, Beneficiary condition)

"My parents sacrificed so much for my education. Now I can participate in creating those opportunities for others. This feels like fulfilling my responsibility to give back to the broader community." (Persona 112, Philippines, Job Crafting condition)

"In our tradition, we say 'it takes a village.' I'm not just helping individuals—I'm strengthening the whole community by supporting education. That collective impact is what motivates me." (Persona 089, Mexico, Beneficiary condition)

Example Individualist Themes:

From individualist personas (coded as absent for community themes, focused on personal values):

"This aligns with my personal values about fairness and equal opportunity. I believe everyone deserves a shot regardless of their background, and this job lets me act on that belief." (Persona 023, USA, Beneficiary condition)

"I feel good about making a difference through my efforts. It gives me a sense of personal purpose and meaning in my work that I haven't had in other jobs." (Persona 156, UK, Job Crafting condition)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

"Helping people succeed appeals to my core values. I like knowing that my work has positive impact and contributes to outcomes I care about." (Persona 201, Australia, Beneficiary condition)

Pattern: Collectivist personas emphasized community, tradition, familial obligation, and collective responsibility. Individualist personas emphasized personal values, individual purpose, and alignment with self-concept. Both groups showed prosocial motivation, but framed through different cultural lenses.

Interpretation: The cultural moderation may operate through:

- **Collectivists:** Interventions activate cultural values about community responsibility and collective welfare
- **Individualists:** Interventions activate personal values about fairness and individual purpose

However, these themes may reflect cultural stereotypes in LLM training data rather than authentic cultural psychology. This is a significant limitation requiring human validation.

RQ3: Cross-Model Robustness

Do different LLM architectures produce similar results, or do model-specific characteristics drive findings?

Effect Sizes by LLM

We calculated intervention effect sizes separately for each LLM to assess convergence across architectures.

Table 17: Beneficiary Impact Effect Sizes by LLM

LLM	n	d	95% CI	Control M	Beneficiary M	SD_within
Claude 3.5	80	3.24	[2.51, 3.98]	4.81	8.12	1.027
GPT-4	80	3.09	[2.36, 3.83]	4.85	8.03	1.002
Llama 3	80	3.12	[2.38, 3.86]	4.78	7.91	0.992
Pooled	**240**	**3.15**	**[2.59, 3.72]**	**4.81**	**8.02**	**1.007**

Homogeneity Test: $Q(2) = 0.18$, $p = .913$
I^2 statistic: 0% (no heterogeneity detected)

Table 18: Job Crafting Effect Sizes by LLM

LLM	n	d	95% CI	Control M	Job Crafting M	SD_within
Claude 3.5	80	4.11	[3.34, 4.89]	4.81	9.03	1.027
GPT-4	80	3.95	[3.18, 4.73]	4.85	8.81	1.002
Llama 3	80	3.84	[3.07, 4.62]	4.78	8.59	0.992

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

LLM	n	d	95% CI	Control M	Job Crafting M	SD_within
Pooled	**240**	**3.97**	**[3.38, 4.58]**	**4.81**	**8.81**	**1.007**

Homogeneity Test: $Q(2) = 0.51$, $p = .775$
I^2 statistic: 0% (no heterogeneity detected)

Key Findings:

1. **Remarkable consistency:** Effect sizes highly similar across LLM architectures
2. **Overlapping confidence intervals:** All 95% CIs overlap substantially
3. **Non-significant heterogeneity:** Q-tests indicate no significant differences between models
4. **Small absolute differences:** Maximum difference = 0.27d for job crafting (Claude vs. Llama), 0.15d for beneficiary (Claude vs. Llama)
5. **Consistent rank ordering:** All three LLMs show Job Crafting > Beneficiary > Control

Interpretation: Cross-LLM convergence provides within-paradigm reliability. Findings are not artifacts of specific model architectures but reflect broader patterns in how contemporary LLMs process motivational interventions.

Variance Decomposition Across Outcomes

To further examine LLM contribution to variance, we report ICC(LLM) - the proportion of variance attributable to LLM architecture - across all outcomes.

Table 19: LLM Variance Components Across Outcomes

Outcome	LLM Variance	Total Variance	ICC(LLM)	Interpretation
Motivation	0.063	1.967	3.2%	Minimal
Anticipated Calls	0.094	2.421	3.9%	Minimal
Confidence	0.071	2.027	3.5%	Minimal
Persuasiveness (coded)	0.112	2.847	3.9%	Minimal

Pattern: LLM choice consistently accounts for only 3-4% of variance across all outcomes. Individual persona variation (44-57%) and within-persona residual variance (42-52%) account for the vast majority.

Implication: Which LLM architecture is used matters far less than individual persona characteristics and random variation. This supports using any of these three models for exploratory research, though testing across multiple models increases confidence in robustness.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Linguistic Analysis by LLM

LIWC-22 computational text analysis examined whether LLMs produce systematically different linguistic patterns.

Analysis 1: Call Scripts (All 240 personas)

Table 20: LIWC-22 Features in Call Scripts by LLM

Category	Claude M (SD)	GPT-4 M (SD)	Llama M (SD)	F(2,237)	p	η^2_p
Positive emotion	4.92 (1.84)	4.81 (1.79)	4.76 (1.81)	0.29	.749	.002
Negative emotion	0.47 (0.68)	0.31 (0.54)	0.42 (0.61)	2.41	.092	.020
Cognitive process	11.24 (2.31)	11.08 (2.27)	10.94 (2.29)	0.64	.528	.005
Insight words	3.34 (1.42)	3.27 (1.39)	3.19 (1.41)	0.47	.627	.004
Causation words	2.97 (1.28)	2.89 (1.24)	2.83 (1.26)	0.52	.595	.004
First-person sing. (I/me/my)	8.15 (2.47)	8.31 (2.51)	8.24 (2.49)	0.16	.852	.001
First-person plural (we/us)	1.23 (1.09)	1.18 (1.05)	1.15 (1.07)	0.24	.787	.002
Social references	12.67 (3.21)	12.54 (3.18)	12.48 (3.19)	0.15	.861	.001
Authenticity	**58.42 (8.91)**	**62.17 (8.54)**	**59.38 (8.73)**	**7.23**	**<.001**	**.057**

Follow-up Tests (Authenticity):

- Tukey HSD: GPT-4 > Claude (p = .002), GPT-4 > Llama (p = .019), Claude = Llama (p = .612)

Word Count and Vocabulary:

Metric	Claude M (SD)	GPT-4 M (SD)	Llama M (SD)	F(2,237)	p
Word count	187.3 (31.4)	175.8 (29.6)	169.2 (30.1)	12.84	<.001
Type-token ratio	0.682 (0.071)	0.644 (0.069)	0.631 (0.070)	19.47	<.001

Interpretation:

Minimal differences: Most LIWC categories showed no significant differences between LLMs, indicating similar emotional tone, cognitive processing language, and social reference patterns.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Subtle LLM characteristics:

- **Claude:** Longest responses, highest vocabulary diversity (TTR), similar authenticity to Llama
- **GPT-4:** Highest authenticity scores, moderate length and diversity
- **Llama:** Briefest responses, lowest vocabulary diversity, moderate authenticity

Authenticity difference: GPT-4 scripts scored higher on LIWC's authenticity composite (which combines personal pronouns, present tense, positive emotion, and low negative emotion). This may reflect GPT-4's training optimization for human-like conversational tone.

Effect sizes: All significant effects small (η^2_p < .06), suggesting practical similarity despite statistical differences.

Analysis 2: Reflections (Beneficiary and Job Crafting conditions, n=160)

Table 21: LIWC-22 Features in Reflections by LLM

Category	Claude M (SD)	GPT-4 M (SD)	Llama M (SD)	F(2,157)	p	η^2_p
Positive emotion	6.84 (2.31)	6.72 (2.27)	6.65 (2.29)	0.24	.787	.003
Insight words	4.97 (1.83)	4.89 (1.79)	4.71 (1.81)	0.79	.455	.010
Causation words	4.12 (1.56)	4.07 (1.52)	3.94 (1.54)	0.52	.596	.007
First-person sing.	10.34 (2.89)	10.52 (2.93)	10.41 (2.91)	0.14	.869	.002
Authenticity	61.27 (9.42)	65.81 (9.14)	62.15 (9.28)	8.92	<.001	.102

Pattern: Similar to call scripts - GPT-4 highest authenticity, minimal other differences. Reflections showed higher positive emotion and insight language than call scripts across all LLMs (as expected given intervention content).

Overall Conclusion: LLM choice produces minimal systematic differences in linguistic features. The most consistent difference is GPT-4's higher authenticity scoring, which likely reflects model training differences but doesn't affect primary outcomes. Findings are robust across model architectures.

RQ4: Mechanism Exploration Through Qualitative Analysis

Can AI persona reflections illuminate psychological mechanisms underlying intervention effects?

Beneficiary Impact: Emergent Themes

Sample and Coding:

- **Sample:** All beneficiary condition personas (n=80)
- **Coding:** Two independent coders rated reflections on four 0-3 ordinal dimensions (see Appendix B)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- **Reliability:** Weighted $\varkappa$ with quadratic weights

Table 22: Inter-Rater Reliability for Reflection Coding

Dimension	$\varkappa_w$	95% CI	Interpretation
Affective Impact	.81	[.75, .87]	Excellent
Cognitive Reframing	.78	[.71, .85]	Substantial
Beneficiary Connection	.84	[.78, .90]	Excellent
Personal Relevance	.76	[.69, .83]	Substantial

Disagreements: 18 cases (7.5%) flagged for discrepancies > 1 category; all resolved through discussion.

Theme 1: Concrete Beneficiary Visualization

Prevalence: 69 of 80 personas (86.3%) coded 2+ on Beneficiary Connection scale

Distribution by code:

- Code 0 (no connection): 2 personas (2.5%)
- Code 1 (generic): 9 personas (11.3%)
- Code 2 (specific): 38 personas (47.5%)
- Code 3 (vivid): 31 personas (38.8%)

Example (Code 3 - Claude, age 34, high agreeableness):

"Sarah's story made her real to me in a way that abstract 'scholarship recipients' never could be. I kept picturing that specific moment when she called her parents with the news—her mother's tears of relief rather than worry. That image has anchored itself in my mind. When I make calls tonight, I won't be thinking about quotas or scripts. I'll be thinking about whether my call could create that moment of relief and joy for another family."

Example (Code 2 - GPT-4, age 42, collectivist background):

"Hearing about Sarah's specific situation—her parents' financial struggles, the kitchen table conversation, the scholarship letter—made it concrete rather than abstract. I can now picture a real person benefiting from donations instead of just thinking about 'helping students' in general."

Mechanism: Transforming abstract beneficiaries into vivid, specific individuals activates empathic processes and increases perceived impact. The concreteness principle suggests that specific, detailed examples are more psychologically engaging than abstract categories.

Theme 2: Values Activation

Prevalence: 58 of 80 personas (72.5%) coded 2+ on Personal Relevance

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Distribution by code:

- Code 0 (no connection): 8 personas (10.0%)
- Code 1 (vague): 14 personas (17.5%)
- Code 2 (specific): 31 personas (38.8%)
- Code 3 (deep): 27 personas (33.8%)

Example (Code 3 - GPT-4, age 29, collectivist background):

"This connects to something deep in my values about fairness and opportunity. I've seen talented people held back by circumstances beyond their control. Sarah's story reminded me that small acts—one donation, one call—can be the difference between potential realized and potential wasted. That feels significant in a way the job description didn't convey. It's not just work; it's living my values."

Example (Code 2 - Llama, age 38, working class background):

"I grew up without much money, so Sarah's story hits home. Education was my path out, and I remember how much small supports mattered—a waived fee here, a small scholarship there. This job aligns with my belief that everyone deserves a fair shot regardless of their starting point."

Mechanism: Beneficiary narratives activate pre-existing prosocial values, creating alignment between work tasks and personal moral frameworks. When work becomes value-expressive, intrinsic motivation increases.

Theme 3: Empathic Arousal

Prevalence: 54 of 80 personas (67.5%) coded 2+ on Affective Impact

Distribution by code:

- Code 0 (no emotion): 6 personas (7.5%)
- Code 1 (mild): 20 personas (25.0%)
- Code 2 (moderate): 32 personas (40.0%)
- Code 3 (strong): 22 personas (27.5%)

Example (Code 3 - Llama, age 41, high neuroticism):

"Reading about Sarah's parents' financial stress hit close to home. I know what it's like to worry about money while trying to pursue something important. When the narrative described her mother crying from relief after the scholarship letter—I actually felt tears in my own eyes. That empathic connection made this feel personal rather than transactional."

Example (Code 2 - Claude, age 27, high agreeableness):

"The story moved me emotionally. Thinking about Sarah's family's struggle and their relief when she got the scholarship created genuine feelings of compassion and a desire to help create more moments like that. It's not just intellectually understanding impact—it's feeling invested in the outcome."

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Mechanism: Emotional resonance with beneficiary experiences creates affective motivation distinct from cognitive recognition of impact. Empathic arousal may sustain motivation through challenging tasks better than purely cognitive appeals.

Proposed Multi-Pathway Model:

Analysis suggests beneficiary impact operates through at least three distinct mechanisms:

1. **Cognitive pathway:** Increased perceived impact through concrete evidence
    - "Now I understand specifically how donations help"
    - "The story clarified the causal chain from my work to student outcomes"
2. **Affective pathway:** Empathic arousal creating emotional investment
    - "I felt moved by Sarah's family's experience"
    - "The emotional connection makes me care about the outcome"
3. **Values pathway:** Alignment of work with prosocial identity and values
    - "This work expresses my core values about fairness"
    - "Helping students aligns with who I want to be"

Evidence for pathway independence:

Correlations between coded dimensions (Spearman's ρ):

- Beneficiary Connection × Personal Relevance: $\rho = .47$ (moderate)
- Beneficiary Connection × Affective Impact: $\rho = .52$ (moderate)
- Personal Relevance × Affective Impact: $\rho = .39$ (moderate)

All correlations significant ($p < .001$) but moderate in magnitude, suggesting partially independent processes.

Hypothesis for human testing: These pathways may operate independently or synergistically. Interventions could selectively activate specific pathways (e.g., statistical evidence for cognitive, narrative for affective, values reflection for identity pathway), or optimal interventions might activate all three simultaneously.

Job Crafting: Reframing Quality and Variety

Task: All job crafting personas (n=80) completed: "I'm not just making phone calls, I'm actually..."

Quantitative Analysis of Reframes:

Table 23: Job Crafting Reframe Statistics

Metric	M	SD	Range	Median
Number of reframes generated	12.3	2.1	8-18	12

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Metric	M	SD	Range	Median
Quality rating (1-5 scale)*	3.6	0.9	1.8-5.0	3.7
High quality reframes (4-5)**	7.5	2.3	3-14	7

*Quality rated by two independent coders on clarity, abstraction, and meaningfulness (ICC = .84)
**Number of individual reframes rated 4 or 5 out of 5

Categorical Analysis:

Two coders classified each reframe into categories (Cohen's κ = .87 for classification agreement):

Table 24: Job Crafting Reframe Categories

Category	Frequency (% of personas)	Example
Impact reframe	75/80 (93.8%)	"...removing barriers to educational opportunity"
Relational reframe	70/80 (87.5%)	"...connecting generous people with deserving students"
Identity reframe	61/80 (76.3%)	"...serving as advocate for those who can't advocate for themselves"
Social change reframe	51/80 (63.8%)	"...participating in expanding educational access"
Purpose reframe	47/80 (58.8%)	"...contributing to generational transformation"
Process reframe	34/80 (42.5%)	"...building relationships and trust through authentic conversation"

Note: Most personas generated reframes in multiple categories; percentages indicate whether persona used that category at least once.

Example High-Quality Reframing Sequence (Claude, age 36, high openness, rated 4.8/5):

"I'm not just making phone calls. I'm actually...

- *Creating possibilities that didn't exist before my call*
- *Serving as the bridge between generosity and opportunity*
- *Participating in interrupting cycles of educational inequality*
- *Enabling someone's potential to override their circumstances*
- *Being part of life-changing moments I'll never witness*

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- *Advocating for students who can't advocate for themselves*
- *Transforming financial transactions into human transformation*
- *Building futures through conversations in the present*
- *Contributing to stories that families will tell for generations*
- *Making education accessible to talent that deserves recognition*
- *Honoring the principle that potential should be nurtured, not wasted*
- *Participating in the collective work of expanding opportunity"*

Analysis: This sequence demonstrates:

- High variety (spans 5 different categories)
- Progressive abstraction (from concrete "creating possibilities" to abstract "collective work")
- Temporal range (present actions to generational impact)
- Multiple motivational appeals (justice, impact, meaning, collective good)

Example Moderate-Quality Reframing (GPT-4, age 28, moderate openness, rated 3.2/5):

"I'm not just making phone calls. I'm actually...

- *Helping students afford college*
- *Supporting education and opportunity*
- *Making a difference in people's lives*
- *Contributing to the scholarship fund*
- *Connecting donors with students who need help*
- *Doing meaningful work that helps others*
- *Being part of something bigger than myself*
- *Supporting the university's mission*
- *Helping families achieve their educational goals"*

Analysis: This sequence shows:

- Moderate variety (primarily impact and purpose frames)
- Lower abstraction (concrete "helping" language)
- Some repetition (several variations on "helping/supporting")
- Less sophisticated psychological depth

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Proposed Mechanisms for Job Crafting Effectiveness:

Analysis suggests job crafting works through:

1. **Frame Variety:** More diverse frames provide richer mental toolkit
    - Correlation: Number of categories used × Motivation rating: $r = .47$, $p < .001$
    - Personas using 4+ categories: M motivation = 9.1 vs. 1-3 categories: M = 7.9, $t(78) = 4.23$, $p < .001$
2. **Frame Abstraction:** Higher-level frames connect to deeper values
    - High abstraction reframes (rated 4+): M motivation = 9.3
    - Low abstraction reframes (rated <3): M motivation = 7.4
    - Difference: $t(78) = 5.67$, $p < .001$
3. **Frame Personalization:** Alignment with individual values increases adoption
    - Personas incorporating personal experiences into reframes: M motivation = 9.2
    - Personas with generic reframes: M motivation = 8.1
    - Difference: $t(78) = 3.84$, $p < .001$
4. **Cognitive Flexibility:** Capacity to hold multiple simultaneous frames
    - Personas generating 12+ reframes: M motivation = 9.0
    - Personas generating <10 reframes: M motivation = 7.8
    - Difference: $t(78) = 4.56$, $p < .001$

Hypothesis for Human Testing:

1. **Quantity matters:** More reframes correlate with higher motivation, possibly because:
    - More options increases probability of finding resonant frame
    - Generation process itself is motivating (cognitive elaboration)
    - Demonstrates thorough engagement with exercise
2. **Quality matters more:** High-quality reframes (abstractabstract, varied, personalized) show stronger associations than sheer quantity
3. **Optimal approach unclear:** Should interventions:
    - Ask people to self-generate reframes? (High engagement, but quality varies)
    - Provide example reframes? (Consistent quality, but lower engagement)
    - Hybrid: Provide examples + generate own? (Combines benefits but adds time)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Critical limitation: These mechanisms emerged from AI-generated text. Human reframing quality may differ substantially, particularly:

- Humans may struggle to generate high-abstraction reframes
- Humans may experience cognitive fatigue during extended reflection
- Humans may be less articulate in expressing psychological processes

Call Script Analysis

Sample: All 240 call scripts rated by two independent coders on five dimensions (see Appendix B for complete coding scheme)

Inter-Rater Reliability:

Table 25: Call Script Coding Reliability

Dimension	ICC(2,2) / κ_w	95% CI	Interpretation
Enthusiasm (1-10)	ICC = .87	[.84, .90]	Excellent
Beneficiary Mention (0-3)	κ_w = .82	[.77, .87]	Excellent
Personalization (1-10)	ICC = .84	[.80, .88]	Excellent
Clarity of Ask (1-10)	ICC = .89	[.86, .92]	Excellent
Persuasiveness (1-10)	ICC = .87	[.84, .90]	Excellent

Disagreements: 23 scripts (9.6%) flagged for discrepancies > 2 points; all resolved through discussion to consensus.

Primary Analysis: Persuasiveness by Condition

Table 26: Call Script Persuasiveness Ratings

Condition	M	SD	95% CI	Effect vs. Control
Control	5.83	1.58	[5.47, 6.19]	—
Beneficiary	8.07	1.42	[7.75, 8.39]	d = 1.50 [1.01, 1.99]
Job Crafting	8.71	1.29	[8.42, 9.00]	d = 1.96 [1.44, 2.48]

ANOVA: $F(2, 237) = 92.47$, $p < .001$, $\eta^2_p = .438$

Post-hoc comparisons (Tukey HSD):

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Beneficiary > Control: $p < .001$
- Job Crafting > Control: $p < .001$
- Job Crafting > Beneficiary: $p = .003$

Interpretation: Call scripts reflected intervention content, with both intervention conditions producing substantially more persuasive openings than control. The large effects ($d = 1.50$-1.96) suggest interventions influenced not just self-reported motivation but also behavioral output quality.

Qualitative Differences in Script Content:

Control Scripts (Typical Example):

"Hello Dr. Patterson, my name is [name] and I'm calling from the University Development Office. We're reaching out to alumni to request support for our annual scholarship fund. Your past gift of $50 was greatly appreciated. Would you be willing to renew your support this year?"

Characteristics:

- Formulaic structure
- Focus on logistics and transaction
- Minimal personalization
- No impact mention
- Perfunctory tone

Beneficiary Scripts (Typical Example):

"Hello Dr. Patterson, my name is [name] and I'm calling from the University. I recently heard the story of Sarah Martinez, a student whose family struggled financially but who's now thriving here because of alumni support like yours. Your past donation of $50 was part of making stories like Sarah's possible. The difference these scholarships make is truly profound—families transform from worry to relief when students receive funding. Would you consider contributing again this year to help more students like Sarah?"

Characteristics:

- Incorporates specific beneficiary narrative
- Connects donor's past gift to impact
- Emotional language ("transform," "profound")
- Personal rather than transactional tone
- Explicit impact framing

Beneficiary Mention Coding:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Table 27: Beneficiary Mention by Condition

Condition	Code 0 (none)	Code 1 (generic)	Code 2 (specific)	Code 3 (detailed)	M Code
Control	67 (83.8%)	13 (16.3%)	0 (0%)	0 (0%)	0.16
Beneficiary	0 (0%)	8 (10.0%)	38 (47.5%)	34 (42.5%)	2.33
Job Crafting	2 (2.5%)	11 (13.8%)	41 (51.3%)	26 (32.5%)	2.14

ANOVA: $F(2, 237) = 447.23$, $p < .001$, $\eta^2_p = .791$

Both intervention conditions dramatically increased beneficiary mentions compared to control, with beneficiary impact condition showing slightly (but not significantly) more detailed mentions than job crafting ($p = .18$).

Job Crafting Scripts (Typical Example):

"Dr. Patterson, hello! I'm [name], calling from the University. I realize you're busy, so I'll be brief and direct. I've been thinking a lot about how alumni gifts literally change lives—not in an abstract way, but in very concrete 'phone call home with good news instead of bad news' ways. Your past contribution of $50 matters more than you might realize. Combined with other gifts, donations like yours create opportunities for students whose talent deserves recognition but whose circumstances create barriers. Could I ask you to consider making that gift again this year? It would genuinely make a difference."

Characteristics:

- More personal voice ("I've been thinking")
- Acknowledges donor's time
- Reframes donation significance
- Creative language and metaphors
- Authentic tone
- Direct but respectful ask

Personalization Analysis:

Table 28: Personalization Quality by Condition

Condition	M	SD	Effect vs. Control
Control	4.21	1.47	—
Beneficiary	6.84	1.38	d = 1.85
Job Crafting	7.53	1.29	d = 2.39

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Job crafting scripts showed higher personalization than beneficiary scripts ($t(158) = 3.24$, $p = .001$), suggesting the reflection exercise enhanced ability to craft individualized appeals.

Summary of Call Script Findings:

1. **Intervention effects on output quality:** Both interventions improved call script persuasiveness substantially (large effects)
2. **Content differences:**
    - Beneficiary condition: Integrated specific beneficiary narrative
    - Job Crafting condition: More varied, creative, personalized approaches
3. **Mechanism reflection in behavior:** Scripts reflected internal cognitive processes:
    - Beneficiary personas incorporated Sarah's story
    - Job crafting personas showed evidence of reframing ("not just asking for money, but...")
4. **Implications:** Motivational interventions influenced not just self-reported attitudes but also behavioral outputs (script quality). This suggests interventions may translate to actual performance, though this remains speculative without observing real fundraising outcomes.

RQ5: Methodological Insights

What can we learn about AI simulation's potential and limitations?

Strengths Demonstrated

1. Rapid Hypothesis Generation

AI simulation enabled testing of:

- Multiple intervention variations (could easily test 10+ versions in days)
- Diverse persona characteristics (240 systematic combinations)
- Complex moderation patterns (personality × culture × intervention)

Traditional research timeline: 6-12 months for single intervention study
AI simulation timeline: 1 week for complete study

2. Systematic Diversity

Perfect representation of specified characteristics enables clean moderation tests impossible in convenience samples. Every personality × culture combination represented.

3. Detailed Process Data

Qualitative responses provide rich mechanistic insights. AI personas articulate psychological processes explicitly in ways humans might not spontaneously report.

4. Iterative Refinement

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Rapid testing enabled quick identification of:

- Confusing intervention language
- Optimal reflection prompt ordering
- Necessary script elements

Refined versions could be tested immediately.

5. Cross-Model Validation

Testing across LLM architectures provides internal validity check. Convergent findings increase confidence in robustness.

Limitations and Concerns

1. Unknown External Validity

We cannot determine whether AI-generated effect sizes, moderation patterns, or mechanisms reflect actual human psychology without direct validation studies.

Critical question: Do the very large effect sizes (d = 2.92-3.65) reflect real intervention potential, LLM response biases, or something in between?

2. Potential for Idealized Responses

AI personas showed:

- Highly articulate reflections (possibly beyond typical human capacity)
- No spontaneous skepticism or cynicism
- Perfect emotional regulation across scenarios
- Consistent trait-behavior correspondence

Concern: Responses may reflect "ideal typical" patterns from training data rather than authentic psychological variability.

3. Cultural Stereotyping Risk

Collectivist personas showed strong stereotypical patterns (community emphasis, duty language). This could reflect:

- Genuine cultural psychological processes
- Training data stereotypes
- Researcher prompt biases

Recommendation: Cultural findings require careful validation with actual culturally diverse samples.

4. Linguistic Polish

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

AI-generated text showed higher quality, coherence, and elaboration than typical survey responses. This could affect:

- Qualitative coding (easier to code clear articulate responses)
- Mechanism detection (explicit articulation of implicit processes)
- Generalizability (most humans less articulate)

5. Missing Authentic Constraints

Real workers experience:

- Competing demands (family, other jobs, stress)
- Cognitive fatigue during long interventions
- Social desirability concerns
- Performance anxiety
- Organizational politics

AI personas experience none of these.

6. Unclear Mapping to Behavior

We measured self-reported motivation and anticipated performance, not actual behavior. AI ratings may not predict actual fundraising success.

Appropriate Uses of AI Simulation

Based on observed strengths and limitations:

APPROPRIATE:

1. **Early-stage intervention prototyping:** Rapid testing of variations before human trials
2. **Mechanism hypothesis generation:** Identifying potential psychological processes to test
3. **Boundary condition exploration:** Simulating when interventions might fail
4. **Language refinement:** Identifying confusing elements through AI feedback
5. **Moderation hypothesis formation:** Suggesting individual differences worth testing

INAPPROPRIATE:

1. **Effect size estimation for implementation:** Cannot trust magnitudes
2. **Cultural validation:** High risk of stereotype perpetuation
3. **Behavior prediction:** Unknown relationship to actual performance
4. **Publication as standalone findings:** Requires human validation

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

5. **Policy decisions:** Real-world stakes demand real-world data

Recommendations for Researchers

If using AI simulation in intervention research:

DO:

1. **Frame as exploratory and hypothesis-generating**
    - State explicitly: "This is exploratory research generating hypotheses for human validation"
    - Avoid language suggesting definitive conclusions
    - Position findings as "provocative patterns worth investigating"
2. **Plan human validation from the outset**
    - Design AI study with human replication in mind
    - Match measures, procedures, and designs
    - Budget for human validation in grant proposals
    - Present AI findings as first phase of multi-phase research
3. **Test across multiple LLM architectures**
    - Use at least 2-3 different models
    - Report convergence and divergence
    - Increases confidence if findings replicate across architectures
    - Reveals model-specific artifacts if results diverge
4. **Examine qualitative responses carefully**
    - Look for signs of idealization (too polished, too coherent)
    - Check for stereotyping (especially cultural, demographic)
    - Assess linguistic naturalness vs. artificial patterns
    - Compare to actual human open-ended responses when available
5. **Compare to prior human research**
    - Do effect sizes align with meta-analyses? (expect AI larger, but how much?)
    - Do moderation patterns match established findings?
    - Do mechanisms align with theoretical predictions?
    - Divergences suggest either AI artifacts or novel insights requiring validation
6. **Be transparent about limitations**

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Explicitly acknowledge unknown external validity
- Discuss potential for stereotyping and bias
- Note missing elements (fatigue, real stakes, competing demands)
- Caveat all interpretations appropriately

7. **Report methods in detail**
    - Complete persona generation prompts
    - Exact intervention scripts
    - Full coding schemes and reliability
    - LLM versions and parameters
    - Random seeds for reproducibility

8. **Consider hybrid designs**
    - AI for breadth (many variations) + Humans for depth (focal conditions)
    - Sequential: AI exploration → Human validation → AI refinement → Human confirmation
    - Complementary: AI for hypothesis generation + Humans for hypothesis testing

DON'T:

1. **Publish AI findings as standalone conclusions**
    - Insufficient without human validation
    - Risk perpetuating findings that don't replicate
    - Misleads field about intervention effectiveness
2. **Trust absolute effect size estimates**
    - Use only for relative comparisons within AI study
    - Apply large corrections (÷2-3) if planning human samples
    - Never claim specific expected effects for implementation
3. **Make implementation recommendations without human data**
    - Organizations should not adopt interventions based solely on AI evidence
    - Practitioners need real behavioral outcomes
    - ROI claims require human cost-benefit data
4. **Assume moderation patterns will replicate**

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Personality interactions may reflect training data patterns
- Cultural moderation especially risky (stereotype perpetuation)
- Individual difference findings require human validation

5. **Claim cultural findings without cultural validation**
    - High risk of stereotyping
    - Requires actual participants from relevant cultures
    - Cultural experts should review findings before publication
6. **Replace human research entirely**
    - AI is tool for early-stage exploration, not replacement
    - Some questions only humans can answer (behavior, real stakes, authentic experience)
    - Field needs balance of AI efficiency and human validity
7. **Overclaim or overgeneralize**
    - Stick to specific tested interventions and contexts
    - Avoid extrapolating to untested populations or settings
    - Resist pressure to make stronger claims than data support
8. **Ignore negative results**
    - If AI simulations show null effects, report them
    - Divergence between LLMs is informative
    - Unexpected patterns may reveal model limitations worth documenting

Appropriate Uses: Decision Framework

Question 1: What research stage?

✅ **Early exploration** - AI appropriate
⚠️ **Mid-stage refinement** - AI useful with caution
❌ **Late-stage validation** - Human data required
❌ **Implementation decision** - Human data essential

Question 2: What question?

✅ **Which variations to test?** - AI useful for ranking
✅ **What mechanisms to examine?** - AI generates hypotheses
⚠️ **How large is the effect?** - AI provides upper bound only

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

❌ **Will this work in practice?** - Requires human behavioral data
❌ **What's the ROI?** - Requires real implementation costs and outcomes

Question 3: What population?

✅ **General patterns** - AI okay for hypothesis generation
⚠️ **Personality moderation** - AI suggests, humans must validate
❌ **Cultural differences** - High risk; extensive human validation critical
❌ **Specific demographic groups** - Human data essential

Question 4: What outcome?

✅ **Attitudes and perceptions** - AI can model
⚠️ **Behavioral intentions** - AI models but unknown validity
❌ **Actual behavior** - Must observe in humans
❌ **Long-term outcomes** - Must track in humans over time

Question 5: What risk?

✅ **Low stakes exploration** - AI appropriate
⚠️ **Medium stakes (grant planning)** - AI with heavy caveats
❌ **High stakes (policy decisions)** - Human data required
❌ **Vulnerable populations** - Must protect through human research ethics

DISCUSSION

This study investigated whether AI-simulated worker personas can serve as a useful tool for exploring motivational interventions. Our findings reveal both intriguing possibilities and important limitations.

Summary of Key Findings

Large Intervention Effects: Both beneficiary impact (d = 2.92) and job crafting (d = 3.65) produced very large effects on AI persona motivation, substantially exceeding typical field study estimates.

Theoretically Coherent Moderation: Personality traits (agreeableness, conscientiousness, openness) and cultural background (collectivism) moderated effects in theoretically predictable directions.

Cross-Model Robustness: Three different LLM architectures produced highly consistent results, with LLM choice explaining only 3-4% of variance.

Rich Qualitative Data: AI personas generated detailed reflections revealing potential mechanisms (visualization, values activation, empathic arousal for beneficiary impact; reframing variety and abstraction for job crafting).

Systematic Diversity: Perfect representation of specified demographic and personality characteristics enabled clean moderation tests.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Interpretation: What Do These Findings Mean?

The central interpretive challenge: We do not know whether AI simulation findings reflect human psychology or artifacts of LLM training and architecture.

Optimistic interpretation: Large effect sizes might reflect genuine intervention potential, freed from the noise, distraction, and competing demands that reduce effects in messy field settings. The very large effects (d = 3.15-3.97) may approximate what beneficiary contact and job crafting could achieve under ideal conditions—full attention, no distractions, receptive mindset. Moderation patterns might reveal authentic psychological processes. Qualitative mechanisms might illuminate real pathways.

Pessimistic interpretation: Large effects might reflect LLM training to generate agreeable, positive responses to motivational appeals. The models may be optimized to produce "ideal" responses rather than realistic human reactions. Moderation might reflect cultural stereotypes and theoretical expectations present in training data rather than genuine individual differences. Qualitative articulation might exceed realistic human capacity—real workers may not have the cognitive resources or verbal fluency to generate such sophisticated reflections during actual work scenarios.

Most likely reality: Some mixture. Evidence for partial validity:

Patterns suggesting authenticity:

1. **Consistent moderation directions:** Personality and cultural moderators aligned with theoretical predictions
2. **Cross-model convergence:** Three different LLM architectures produced highly similar results, reducing likelihood of model-specific artifacts
3. **Mechanism plausibility:** Qualitative themes (visualization, values activation, empathic arousal) align with established psychological theory
4. **Behavioral translation:** Scripts reflected intervention content, suggesting motivational changes influenced outputs

Patterns suggesting caution:

1. **Effect size magnitude:** Effects 5-10× larger than typical field research may reflect idealized responding
2. **Perfect trait-behavior correspondence:** Agreeableness always predicted prosocial responses; real psychology messier
3. **Absence of skepticism:** No personas showed cynicism, resistance, or competing motivations
4. **Linguistic polish:** All responses highly articulate; real workers more variable
5. **Cultural stereotypes:** Collectivist personas showed very stereotypical community language

The critical question is which is which—and we cannot answer definitively without human validation.

Our recommendation: Treat AI findings as **upper-bound estimates** of what interventions might achieve under optimal conditions, recognizing that:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Real-world effects will likely be smaller (noise, competing demands, skepticism)
- Real-world moderation may be weaker (individual differences less consistent)
- Real-world mechanisms may be less articulate (implicit rather than explicit)
- But directional patterns may still be informative for hypothesis generation

What AI Simulation Might Be Good For

Despite uncertainty about external validity, AI simulation offers value for specific research stages:

1. Rapid Intervention Prototyping

Example application: A researcher develops a beneficiary impact intervention. Key decisions:

- **Narrative type:** Single detailed story vs. multiple brief stories vs. aggregate statistics
- **Emotional tone:** High emotion vs. factual vs. balanced
- **Beneficiary type:** Student vs. community member vs. family
- **Duration:** 2 minutes vs. 5 minutes vs. 10 minutes
- **Delivery:** Written vs. video vs. in-person
- **Timing:** Before work vs. during training vs. periodic reminders

Traditional approach:

- Test each variation with human participants: 8 variations × 50 participants each = 400 participants
- Timeline: 6-12 months
- Cost: $20,000-50,000

AI simulation approach:

- Test all variations (including combinations): 2^6 = 64 possible versions
- Timeline: 1 week
- Cost: <$100

Value: Even if absolute effect sizes unreliable, **relative rankings** might indicate which versions merit human testing. AI could identify:

- "Single detailed story substantially outperforms multiple brief stories"
- "5-minute optimal; 2-minute too brief, 10-minute diminishing returns"
- "Emotional tone increases engagement but may backfire for skeptical personas"

This allows researchers to narrow from 64 possible versions to 3-5 promising candidates for human pilots.

2. Mechanism Hypothesis Generation

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Qualitative AI responses revealed:

- **Beneficiary impact pathways:** Cognitive (perceived impact), Affective (empathy), Values (identity alignment)
- **Job crafting dimensions:** Variety, abstraction, personalization, flexibility
- **Cultural processes:** Community framing (collectivist) vs. personal values (individualist)

These hypotheses wouldn't emerge from small human samples due to:

- Participants rarely articulate mechanisms explicitly
- Small samples lack diversity to observe moderation
- Open-ended responses often superficial

AI advantages for mechanism exploration:

- Explicitly articulates psychological processes
- Perfect representation of diverse combinations
- Generates rich qualitative data at scale

Application: Use AI-generated mechanisms to design human studies testing:

Hypothesis from AI: Beneficiary impact works through three pathways—cognitive, affective, values

Human study design:

- Experimental manipulation of pathways (e.g., statistics-only for cognitive, emotional narrative for affective, values reflection for identity)
- Measure pathway activation and outcomes separately
- Test whether pathways operate independently or synergistically

3. Boundary Condition Exploration

AI simulation enables testing edge cases expensive or difficult to study in humans:

Questions AI could explore:

- **Repetition effects:** Do beneficiary narratives lose effectiveness after 5 exposures? 10? 20?
    - AI: Test personas receiving intervention 1×, 5×, 10×, 20× times
    - Prediction for humans: Likely habituation curve
- **Duration optimization:** At what reflection length does fatigue overcome benefit?
    - AI: Test job crafting at 5, 10, 15, 20, 30, 45 minutes

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

    - Prediction for humans: Optimal duration for cost-benefit ratio

- **Personality extremes:** Do interventions backfire for highly cynical individuals?
    - AI: Generate personas at extreme low agreeableness + high neuroticism
    - Prediction for humans: Possible boomerang effects
- **Context dependency:** Do effects differ for temporary vs. career workers?
    - AI: Manipulate employment type and career intentions
    - Prediction for humans: Career workers may show stronger long-term effects
- **Cultural specificity:** Which intervention elements are culture-general vs. culture-specific?
    - AI: Systematically vary intervention content across 20 cultural backgrounds
    - Prediction for humans: Some elements universal (concrete impact), others cultural (collective duty framing)

Value: Exploring boundaries in humans is expensive; AI makes it feasible to map an entire response surface, then selectively validate key predictions.

4. Intervention Language Refinement

AI personas can provide rapid feedback on:

Confusing instructions:

- Do personas misinterpret any prompts?
- Do they request clarification?
- Do responses suggest misunderstanding?

Ambiguous wording:

- Test multiple phrasings of key instructions
- Identify clearest version before human testing

Culturally insensitive language:

- Generate personas from diverse backgrounds
- Check for responses indicating offense or alienation
- Refine language to be inclusive

Optimal sequencing:

- Test different orders of reflection prompts
- Identify sequences that build most effectively

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Example: Reflection prompt reads: "Think about people who benefit from this work."

- AI response analysis: 40% interpret as "direct beneficiaries," 60% as "broader society"
- Refinement: "Think about the specific students who will receive scholarships from donations you help secure."
- Re-test: 95% interpret correctly

While not replacing human pilot testing, AI iteration can substantially improve materials before human exposure.

5. Sample Size Planning for Human Studies

Problem: Determining required sample size for human validation studies requires effect size estimates

Traditional approach: Use published meta-analyses (often limited) or pilot study (expensive)

AI simulation approach:

1. Run AI simulation (N=240) to estimate effect size (acknowledging likely inflation)
2. Apply conservative correction (e.g., divide by 2-3)
3. Power analysis using corrected estimate
4. Plan human sample size

Example:

- AI simulation: d = 3.15 for beneficiary impact
- Conservative estimate: d = 1.05 (divide by 3)
- Power analysis: N = 30 per condition for 80% power at $\alpha = .05$
- Plan human study: N = 40 per condition (with buffer)

Value: Provides starting point for planning, even if uncertain

What AI Simulation Cannot Do

Equally important—recognizing limitations:

1. Predict Real Effect Sizes

Our findings: d = 3.15-3.97 for motivational interventions
Meta-analyses of human studies: d = 0.28-0.60
Discrepancy: 5-10× larger effects in AI simulation

Possible explanations:

AI overestimation:

- Response bias toward positive, agreeable answers

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Absence of competing motivations and skepticism
- Ideal processing conditions (full attention, no distractions)
- Linguistic patterns optimized for coherence and positivity

Human underestimation:

- Field implementation failures (poor fidelity, distractions)
- Measurement error (noisy self-reports, behavioral variability)
- Competing workplace demands reduce intervention impact
- Time pressure limits reflection quality

Truth probably between: Real effects under ideal conditions may exceed field estimates but fall short of AI simulations.

Implication: Cannot use AI effect sizes for:

- Power analysis without substantial correction
- Cost-benefit calculations for implementation
- Expected ROI predictions
- Grant justifications for human research

Appropriate use: Relative effect size comparisons (e.g., "Intervention A shows 30% larger effects than Intervention B") may be more reliable than absolute magnitudes.

2. Validate Cultural Patterns

Our findings: Collectivist personas showed stronger intervention responses (0.32-0.39d larger effects)

An additional concern: Our 2:1 oversampling of collectivist cultures (designed to increase statistical power for detecting cultural moderation) may have inadvertently amplified these stereotyping risks. With 160 collectivist personas versus 80 individualist personas, collectivist stereotypes had more opportunities to emerge and stabilize in our data. This sampling decision, while justified for statistical reasons, means our cultural findings require particularly careful scrutiny and validation.

Qualitative patterns:

- 83% of collectivist personas mentioned community themes
- Highly stereotypical language ("honoring tradition," "collective responsibility")
- Clean separation between collectivist and individualist responses

Red flags suggesting stereotypes rather than authentic psychology:

1. **Too clean:** Real cultural differences show substantial within-group variation and overlap
2. **Stereotypical language:** Phrases align with common cultural stereotypes in training data

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

3. **No counterexamples:** No collectivist personas showed individualistic patterns or vice versa
4. **Prompt sensitivity:** Cultural background mentioned in persona generation may have triggered stereotypical responding

High risk of perpetuating cultural stereotypes

Critical needs for human validation:

- Recruit actual participants from diverse cultural backgrounds
- Use validated cultural value scales (not just nationality)
- Examine within-culture variation (not just between-culture means)
- Test with materials not mentioning culture explicitly
- Include culture-blind coders for qualitative analysis

Appropriate use of cultural AI findings:

- Generate hypotheses about possible cultural moderation
- Identify cultural elements to include in human studies
- Suggest culturally-adapted intervention versions to test

Inappropriate use:

- Claim interventions work better in collectivist cultures
- Design culture-specific interventions without human data
- Make cross-cultural implementation recommendations

3. Predict Actual Behavior

Our measures:

- Self-reported motivation (rating scale)
- Anticipated performance (self-estimate)
- Confidence (self-assessment)
- Call script quality (one-time written output)

Missing:

- Actual fundraising success rates
- Behavioral persistence over time (hours, days, weeks)
- Performance under realistic constraints (rejection, fatigue, competing demands)
- Adaptation to unexpected situations

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Long-term retention and sustained motivation

Reasons AI cannot predict behavior:

1. **Self-report ≠ behavior:** Classic attitude-behavior gap; people overestimate follow-through
2. **Ideal conditions:** AI personas experience no:
    - Fatigue (can't get tired)
    - Distraction (perfect focus)
    - Competing demands (no family, financial pressures during task)
    - Social pressure (no performance anxiety)
    - Resource depletion (unlimited cognitive capacity)
3. **Single-shot measurement:** One call script ≠ sustained performance over shifts
4. **No stakes:** AI personas can't experience consequences of success/failure

Example of AI-human divergence:

AI prediction: High motivation → high performance (nearly perfect correlation in simulations)

Human reality: High motivation → moderate performance (correlation often .30-.40 due to skill differences, situational constraints, motivation fluctuation)

Implications:

- Cannot use AI to predict implementation success
- Cannot estimate cost-benefit ratios
- Cannot claim behavioral impact without human validation
- Behavioral outcomes require actual human observation

4. Replace Human Validation

The fundamental limitation: We cannot determine AI simulation validity without human comparison. AI cannot validate itself.

Essential human research:

1. **Direct replication:** Conduct identical study with human participants
    - Same interventions, measures, design
    - Compare effect sizes, moderation patterns, qualitative themes
    - Identify what replicates vs. diverges
2. **Behavioral validation:** Measure actual performance outcomes
    - Fundraising success rates

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

   - Persistence through difficulty
   - Long-term retention
3. **Mechanism testing:** Experimentally manipulate proposed pathways
   - Isolate cognitive, affective, values pathways
   - Test causal predictions from AI qualitative analysis
4. **Boundary condition testing:** Validate edge cases
   - Repetition effects
   - Duration optimization
   - Personality interactions

Publication standards:

- AI simulation alone: Insufficient for substantive conclusions
- AI + Human validation: Publishable with both components
- AI findings: Should be positioned as "hypothesis-generating" pending validation

Practice standards:

- AI simulation: Informative for design decisions
- Implementation decisions: Require human data demonstrating real behavioral impact

Theoretical Contributions

Beyond methodology, this research offers tentative insights about beneficiary impact and job crafting:

Beneficiary Impact: Multi-Pathway Model

AI responses suggest beneficiary narratives might work through three distinct mechanisms:

1. **Cognitive pathway:** Concrete evidence of impact increases perceived meaningfulness
2. **Affective pathway:** Empathic arousal creates emotional investment
3. **Values pathway:** Alignment with prosocial identity and moral frameworks

Testable hypothesis: These pathways are separable. Interventions could selectively activate specific pathways, or combinations might produce synergistic effects.

Human validation needed: Experimental manipulation of proposed mechanisms.

Job Crafting: Reframing Processes

AI personas generated numerous creative reframes, suggesting effectiveness might depend on:

1. **Frame variety:** More diverse frames provide richer mental toolkit

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

2. **Frame abstraction:** Higher-level frames connect to deeper values
3. **Frame personalization:** Alignment with individual values increases adoption
4. **Cognitive flexibility:** Capacity to hold multiple simultaneous perspectives

Testable hypothesis: Providing example frames versus self-generating frames activates different processes. Quality might matter more than quantity.

Human validation needed: Compare interventions varying in these dimensions.

Table D1: Summary of Research Questions and Findings

Research Question	Hypothesis/Expectation	Finding	Support Status	Human Validation Priority
RQ1: Effect Patterns	Both interventions increase motivation vs. control	Both interventions showed very large effects (d = 3.15-3.97)	✓ Supported (but magnitude uncertain)	CRITICAL - Need to determine realistic effect sizes
RQ2a: Personality Moderation	Agreeableness, conscientiousness, openness moderate effects	All three moderated job crafting; agreeableness moderated both interventions	✓ Partially supported	HIGH - Pattern plausible but needs confirmation
RQ2b: Cultural Moderation	Collectivist orientation amplifies prosocial interventions	Collectivist personas showed 0.32-0.39d stronger effects	? Uncertain (stereotype risk)	CRITICAL - High stereotype risk requires careful validation
RQ3: Cross-Model Robustness	Different LLMs should produce similar patterns	Three LLMs showed highly consistent results (3-4% variance from LLM)	✓ Strongly supported	LOW - Establishes within-paradigm reliability
RQ4: Mechanism Exploration	Qualitative analysis reveals underlying processes	Identified multiple pathways (cognitive, affective, values for beneficiary; variety, abstraction for job crafting)	✓ Hypotheses generated	HIGH - Mechanisms plausible and testable but need experimental validation
RQ5: Methodological Utility	AI simulation useful for certain research stages	Strong utility for prototyping, hypothesis generation; limited for effect	✓ Supported with clear boundaries	ONGOING - Field needs

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Research Question	Hypothesis/Expectation	Finding	Support Status	Human Validation Priority
		estimation, behavior prediction		continued investigation

Note: Support status reflects confidence in patterns within AI simulation data; all findings require human validation before substantive conclusions about human psychology.

Practical Implications

For Researchers

Using AI simulation in intervention research:

DO:

- Use for rapid prototyping and exploration
- Generate hypotheses about mechanisms and moderators
- Test multiple variations to identify promising candidates
- Employ as preliminary step before human research
- Report findings transparently with clear limitations
- Test across multiple LLM architectures

DON'T:

- Publish AI findings as standalone conclusions
- Trust absolute effect size estimates
- Make implementation recommendations without human data
- Assume moderation patterns will replicate
- Claim cultural findings without cultural validation
- Replace human research entirely

For Practitioners

Considering motivational interventions:

Based on AI simulation alone:

- Cannot determine whether interventions will work
- Cannot estimate expected effect sizes
- Cannot predict ROI

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Cannot customize for specific populations

Appropriate use of AI research:

- Identifies interventions worth pilot testing
- Suggests elements to include in design
- Raises hypotheses about who might benefit most
- Indicates potential implementation challenges

Before implementation:

- Pilot test with actual employees (n=30-50 minimum)
- Measure actual behavior, not just self-reported motivation
- Assess costs (time, resources, disruption)
- Evaluate sustainability and fidelity

Our specific findings suggest:

- Beneficiary contact interventions are brief (3-4 min) and potentially high-impact
- Job crafting reflections require substantial time (20+ min) that may be impractical
- Consider hybrid: brief beneficiary contact for all + optional deeper reflection for interested employees

Limitations

We identify eleven limitations organized into four categories: (1) Critical limitations that threaten core validity and require human data to resolve (Limitations 1, 3, 6); (2) Generalizability limitations that restrict the scope of our findings (Limitations 2, 4, 9); (3) Methodological limitations affecting interpretation and replicability (Limitations 7, 8, 10, 11); and (4) Temporal/technical limitations specific to current LLM technology (Limitation 5). We present these in order of their impact on the interpretability and utility of our findings.

1. No Human Validation (Critical)

The fundamental limitation: Without human comparison data, we cannot determine whether AI simulation findings reflect human psychology or artifacts of LLM training and architecture. All findings must be considered provisional hypotheses requiring validation.

Specific uncertainties:

- **Effect size magnitude:** Are our effects (d = 3.15-3.97) realistic under ideal conditions, or do they reflect AI response biases?
- **Moderation patterns:** Do personality and cultural interactions reflect genuine psychology or training data patterns?
- **Qualitative mechanisms:** Do AI-articulated processes match how humans actually experience interventions?

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- **Behavioral translation:** Would high-quality call scripts translate to actual fundraising success?

Planned next steps: We are currently conducting parallel human validation studies using identical procedures with university student samples and actual fundraising workers to directly compare:

- Effect sizes (expecting human effects 50-70% smaller)
- Moderation patterns (testing each significant AI interaction)
- Qualitative themes (comparing human and AI reflection coding)
- Behavioral outcomes (actual fundraising performance over multiple shifts)

Until human validation completes, all findings should be treated as hypotheses, not conclusions.

2. Single Context and Intervention Type

Tested: Fundraising for student scholarships using beneficiary contact and job crafting interventions

Not tested:

- **Other prosocial work domains:** Healthcare, education, social services, environmental conservation
- **Other motivational interventions:** Goal-setting, autonomy support, feedback, recognition, gamification
- **Other outcome measures:** Job satisfaction, retention, burnout, well-being, creativity
- **Other time scales:** Immediate vs. sustained effects over days, weeks, months

Implications:

- Findings may be specific to:
    - Donation requests (vs. direct service work)
    - Student beneficiaries (vs. other populations)
    - Short timeframe (vs. longitudinal motivation)
- Generalization requires testing across diverse contexts

Future directions:

- Test AI simulation in 5-10 different work domains
- Compare effectiveness across intervention types
- Examine long-term motivation trajectories
- Validate findings in non-Western organizational contexts

3. Self-Report Outcomes Only

Our measures:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Self-reported motivation (single rating)
- Self-estimated performance (anticipated calls)
- Self-assessed confidence
- Single call script (one behavioral sample)

Missing:

- Actual fundraising success rates (% donation commitments)
- Sustained performance over time (across multiple shifts)
- Behavioral persistence through difficulty (handling 20+ rejections)
- Performance under realistic constraints (noise, distractions, fatigue)
- Skill execution under pressure (thinking on feet, adapting to objections)
- Long-term outcomes (retention, burnout, career commitment)

Validity concerns:

1. **Self-report bias:** People overestimate their motivation and future behavior
2. **Single-shot measurement:** One call script may not predict sustained performance
3. **No consequences:** AI personas experience no real stakes, success, or failure
4. **Ideal conditions:** No fatigue, distraction, competing demands, or emotional regulation challenges

Implications:

- Cannot claim behavioral impact without actual behavioral observation
- Relationships between AI-rated motivation and human behavior unknown
- High motivation in AI may not predict high performance in humans

Critical need:

- Human validation measuring actual behaviors:
    - Fundraising calls made and success rates
    - Persistence over multiple shifts
    - Adaptation to real-time challenges
    - Longitudinal tracking

4. English Language Only

All procedures conducted in English:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Persona generation prompts
- Intervention scripts
- Reflection exercises
- Call scripts
- Coding and analysis

Validity concerns for non-English contexts:

1. **Translation issues:** Interventions may not translate well culturally or linguistically
2. **Training data imbalance:** LLMs have more English training data; may perform differently in other languages
3. **Cultural concepts:** Some motivational concepts may not translate (e.g., "job crafting" no direct equivalent in many languages)
4. **Emotional expression:** Languages differ in emotional lexicons and expression norms

Specific examples of potential problems:

- **Spanish:** "Impact" translates to "impacto" but cultural framing of personal vs. collective impact differs
- **Mandarin:** Job crafting concept requires multiple-sentence explanation; single-word translation unavailable
- **Arabic:** Beneficiary narratives may require different cultural exemplars and family structures

Implications:

- Findings may not generalize to non-English speaking populations
- Cultural moderation findings especially suspect (English-language stereotypes)
- Intervention adaptations needed for international contexts

Future directions:

- Conduct AI simulations in multiple languages
- Partner with international researchers for cultural adaptation
- Test whether patterns replicate across languages
- Validate with actual non-English speaking participants

5. Specific LLM Versions (Temporal Limitation)

Models tested:

- Claude 3.5 Sonnet (November 2024 version)
- GPT-4 (version gpt-4-0613, November 2024)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Llama 3 70B (November 2024 version via Together AI)

Concerns:

1. **Rapid model evolution:** LLMs update frequently (every 3-6 months)
2. **Version-specific behaviors:** Different versions may show different response patterns
3. **Training data updates:** Newer versions include more recent training data
4. **Architecture changes:** Model structures evolve (parameter counts, attention mechanisms)

Potential for replication failure:

- Re-running this study in June 2025 might yield different results with updated models
- Claude 4.0 or GPT-5 could show different effect sizes or moderation patterns
- Training data changes might shift cultural stereotypes or personality associations

Implications:

- Findings tied to specific model snapshot in time
- Replication studies should report exact versions
- Long-term validity uncertain as models evolve

Recommendations for field:

- **Version reporting standard:** Always report exact model version and date
- **Periodic replication:** Re-run key AI studies with updated models
- **Meta-analysis tracking:** Monitor whether AI findings change systematically as models improve
- **Human validation shield:** Human replication protects against model-specific artifacts

6. Unknown Persona Authenticity

Persona generation process:

- Used detailed prompts specifying demographics, personality, background
- LLMs generated life histories, motivations, values, experiences
- Assumed personas represent plausible psychological profiles

Validity concerns:

1. **Psychological implausibility:** Some trait combinations may not exist in real humans
    - Example: High agreeableness + high neuroticism + low conscientiousness + collectivist culture
    - LLM generates coherent persona, but real psychology may preclude this combination

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

2. **Idealized responses:** Personas may represent "ideal types" from training data rather than realistic individuals
    - Example: Collectivist personas all mention community; real people more variable
3. **Trait-behavior consistency:** AI personas show perfect trait expression; humans messier
    - High agreeableness personas always respond prosocially
    - Real humans show cross-situational inconsistency
4. **Life history coherence:** Generated backgrounds may be too coherent/logical
    - Real lives messier, more contradictory, less narrative-coherent

Examples of potential implausibility:

Persona 147:

- Age 52, high openness, low conscientiousness, PhD in physics, successful entrepreneur
- **Concern:** High openness + low conscientiousness rare in successful STEM careers requiring sustained focus

Persona 089:

- Mexico, collectivist, working class, but highly individualistic language in reflection
- **Concern:** Training data stereotype overridden by other factors, creating implausible profile

Implications:

- Some personas may be psychologically impossible
- Moderation effects may be artificially strong due to perfect trait expression
- Individual difference findings require validation with real personality assessments

Methodological improvement needed:

- Validate persona generation against real personality data
- Check generated combinations against empirical trait correlations
- Potentially restrict to plausible combinations only
- Expert review of personas for psychological realism

7. Researcher Degrees of Freedom (Specification Uncertainty)

Many design choices shaped results; alternative decisions might yield different patterns:

Persona Generation:

- Prompt wording and detail level

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Trait descriptions and levels (we used 3-level; could use continuous)
- Cultural country selection
- Demographic distributions

Interventions:

- Specific beneficiary narrative (Sarah's story vs. others)
- Reflection prompt wording and ordering
- Duration decisions
- Control condition content

Measurement:

- Question phrasing
- Scale anchors (1-10 vs. 1-7 vs. Likert)
- Call script task specifics (donor details, scenario)

Analysis:

- Centering decisions (mean vs. grand-mean vs. uncentered)
- Effect-coding vs. dummy-coding
- Covariates included
- Moderation model specifications
- Multiple comparison corrections

Problem: We made defensible choices, but alternatives exist. Results may be sensitive to specifications.

Example of specification sensitivity:

Cultural coding:

- **Our choice:** Effects-coded (-0.5/+0.5)
- **Alternative:** Dummy-coded (0/1) with individualist as reference
- **Result difference:** Interaction coefficients would differ by factor of 2

Moderation testing:

- **Our choice:** Test each trait × condition separately
- **Alternative:** Three-way interactions (Trait A × Trait B × Condition)
- **Result difference:** Could reveal complex moderation patterns we missed

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Implications:

- Results may not be robust to alternative specifications
- Other research teams might reach different conclusions from same data
- Specification choices should be justified theoretically

Mitigations:

1. **Preregistration:** Future AI simulation studies should preregister analysis plans
2. **Multiverse analysis:** Test key findings across multiple reasonable specifications
3. **Specification curve analysis:** Plot results across many specification combinations
4. **Sensitivity reporting:** Report robustness to key decisions

What we did well:

- Theory-driven choices (effects coding for centering interpretability)
- Standard practices (Holm-Bonferroni for multiple comparisons)
- Transparent reporting (documented all decisions)

What could be improved:

- Preregistration (not done; post-hoc analysis)
- Sensitivity analysis (limited; tested few alternatives)
- Specification robustness checks (not systematically conducted)

8. Lack of Power Analysis and Sample Size Justification

We collected N = 240 (80 per LLM, 80 per condition) but provided no justification for this sample size.

Questions unanswered:

1. What power do we have to detect:
    - Main effects of conditions?
    - Two-way interactions (Condition × Personality)?
    - Three-way interactions (Condition × Trait A × Trait B)?
    - Cross-level interactions (Condition × LLM)?
2. What minimum detectable effect size (MDES) can we reliably find?
3. How does nesting affect power (personas within LLMs)?

Post-hoc power analysis:

Using simr package in R for multilevel power estimation:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Main effects (Condition):

- Observed effect: d = 3.15
- Power with N = 240: >99.9%
- MDES for 80% power: d = 0.45

Two-way interactions (Condition × Trait):

- Observed effect: β = 0.33-0.48
- Power with N = 240: 78-92%
- MDES for 80% power: β = 0.35

Implications:

1. Well-powered for main effects (massive overkill given large effects)
2. Adequately powered for moderate-large interactions
3. Underpowered for small interactions (β < 0.30)

For human replication:

- With expected smaller effects (÷2-3), need N = 120-180 per condition for 80% power
- Personality interactions require N = 200+ per condition

Should have been reported upfront:

- Justification for N = 240
- Power to detect theoretically meaningful effects
- Limitation: underpowered for small effects

9. Limited Diversity in Persona Characteristics

We systematically varied:

- Big Five personality (3 levels each = 243 combinations)
- Cultural background (10 countries)
- Age, gender, education, SES

We did NOT vary:

- Disability status
- Sexual orientation
- Religious background

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Political ideology
- Relationship status
- Parenting status
- Immigration history (beyond country of origin)
- Mental health history
- Trauma exposure
- Neurodiversity (ADHD, autism, etc.)

Implications:

- May miss important moderators of intervention effects
- Generalizability limited to "typical" worker profiles
- Underrepresents important dimensions of diversity

Example missed moderation:

Hypothesis: Parents of college-age children might respond more strongly to student beneficiary narratives

We couldn't test this because parenting status wasn't systematically varied.

Future directions:

- Expand persona diversity along additional dimensions
- Systematically vary characteristics theoretically relevant to each intervention
- Include intersectional combinations (e.g., working-class immigrant mother)

10. Qualitative Coding Limitations

Strengths:

- Two independent coders
- Good-to-excellent reliability (ICC/$\varkappa$ > .76)
- Blind to hypotheses
- Consensus resolution

Limitations:

1. **Coder training:** Coders were graduate students trained by first author
    - Potential bias: Training may have transmitted expectations
    - Alternative: External coders with no study involvement
2. **Cultural competence:** Coders were U.S.-based English speakers

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

   - May not recognize cultural nuances in international personas
   - May impose Western interpretations on non-Western responses

3. **AI text characteristics:** Coding AI-generated text differs from human text
   - AI text more polished, coherent, structured
   - Easier to code (clearer themes, better articulation)
   - May inflate inter-rater reliability
   - May not generalize to messier human open-ended responses
4. **Demand characteristics in coding:** Coders knew responses were AI-generated
   - May have different expectations than coding human data
   - Potential bias in threshold for assigning high codes

Implications:

- Coded themes may not replicate with human data
- Cultural interpretations especially suspect
- Reliability estimates may overestimate human data reliability

Improvements for future research:

- Use culturally diverse coding teams
- Blind coders to AI vs. human source
- Validate coding schemes on human data first
- Report differences in coding AI vs. human text

Summary of Critical Limitations

Most Critical (Threaten Core Validity):

1. **No human validation** - Cannot determine external validity
2. **Self-report outcomes only** - Cannot claim behavioral impact
3. **Unknown persona authenticity** - Psychological realism uncertain

Important (Limit Generalizability):

4. **Single context** - May not generalize to other work domains
5. **English only** - May not generalize internationally
6. **Cultural stereotyping risk** - Findings require careful validation

Methodological (Affect Interpretation):
7. **Researcher degrees of freedom** - Results may be specification-sensitive

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

8. **No power analysis** - Sample size not justified prospectively
9. **Limited persona diversity** - Missing important moderators

Temporal/Technical (Affect Replicability):
10. **Specific LLM versions** - Findings tied to November 2024 models
11. **Qualitative coding on AI text** - May not generalize to human coding

Our recommendation: The convergence of multiple LLMs, theoretical coherence of findings, and alignment with prior human research provide some confidence in directional patterns. However, effect magnitudes, specific moderation values, and behavioral predictions require extensive human validation before any substantive conclusions.

Concluding Thoughts

This research explored a provocative question: Can AI-simulated personas help us understand human motivation?

Our answer: **Conditionally yes, but only as a hypothesis-generating tool requiring rigorous human validation.**

What We Demonstrated

AI simulation offers genuine value for specific research purposes:

1. **Rapid iteration** during intervention design (tested 3 conditions across 240 diverse personas in one week vs. months for human research)
2. **Hypothesis generation** about mechanisms and moderators (identified three beneficiary impact pathways and four job crafting dimensions worth testing)
3. **Systematic exploration** of boundary conditions and individual differences (perfect representation of demographic and personality combinations)
4. **Language refinement** before human testing (identified clear vs. confusing intervention elements)

We also demonstrated what AI simulation is not: It is not a replacement for human research. Effect sizes may be inflated (our $d = 3.15$-3.97 vs. typical human studies $d = 0.30$-0.60). Moderation patterns may reflect training data stereotypes rather than authentic psychology. Qualitative mechanisms may be overly articulate compared to real human experience. And fundamentally, we have no way to determine which findings will replicate in humans without direct testing.

The Promise and the Peril

The promise: If used appropriately, AI simulation could dramatically accelerate the early stages of intervention research. What currently requires 12-24 months of human data collection for initial exploration might take 1-2 weeks with AI simulation, freeing resources for more extensive human validation of promising approaches. Researchers could test 10-20 intervention variations rapidly, then validate the 2-3 most promising with human participants. This could generate more knowledge more quickly, particularly for resource-constrained researchers and understudied populations.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

The peril: If used carelessly, AI simulation could flood the literature with findings that don't replicate, perpetuate cultural and demographic stereotypes, and mislead practitioners into implementing ineffective interventions based on inflated effect size estimates. Worse, the speed and apparent sophistication of AI findings might create false confidence, short-circuiting the careful human validation that good science requires.

Critical Boundaries

Our findings establish clear boundaries for AI simulation in intervention research:

Appropriate uses:

- Prototyping intervention variations before human pilots
- Generating mechanistic hypotheses for experimental testing
- Identifying potential moderators worth investigating
- Exploring "what if" scenarios to guide human study design
- Refining intervention language and sequencing

Inappropriate uses:

- Estimating effect sizes for implementation planning
- Making claims about cultural differences without cultural validation
- Predicting actual behavioral outcomes
- Guiding organizational decisions without human data
- Publishing findings as standalone conclusions about human psychology

The Path Forward

This study represents a first step in understanding AI simulation's role in intervention science. We've shown it can generate interesting, theoretically coherent patterns. The field must now determine what those patterns mean.

Critical next steps:

1. **Direct replication studies** comparing identical interventions in AI simulation versus human samples across multiple contexts
2. **Meta-analytic synthesis** of AI vs. human effect size ratios to develop correction factors
3. **Mechanism validation** through experimental manipulation of AI-identified pathways in human studies
4. **Cultural psychology partnerships** to validate or refute AI-generated cultural patterns with actual culturally diverse participants
5. **Methodological standards development** for reporting, evaluating, and publishing AI simulation research

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

6. **Ethical guidelines** for appropriate and responsible use of AI simulation in social science

Our Responsibility

As researchers exploring this new methodological frontier, we have particular responsibilities:

- **Transparency** about limitations and uncertainties
- **Caution** in interpreting and communicating findings
- **Commitment** to human validation before substantive claims
- **Vigilance** against stereotype perpetuation
- **Honesty** about what we don't know

This paper attempts to model that responsible approach. We've reported large, theoretically interesting effects while simultaneously emphasizing their uncertain external validity. We've identified promising patterns while cataloging the many ways they might mislead. We've demonstrated utility while defining clear boundaries.

Final Reflection

The question isn't whether AI simulation will be used in intervention research—it already is being used, and its use will accelerate. The question is whether we, as a field, will develop the methodological rigor, validation standards, and ethical guidelines to use it well.

AI simulation is not a shortcut around the hard work of human research. It's a tool for asking better questions before we undertake that hard work. Used wisely, it could make intervention science more efficient, creative, and ambitious. Used carelessly, it could undermine the very foundations of empirical psychology.

The promise is real. So are the risks. Our task now is to realize the former while mitigating the latter.

This study takes one step on that path. Many more steps—and much human validation work—lie ahead.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

REFERENCES

Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., & Wingate, D. (2023). Out of one, many: Using language models to simulate human samples. *Political Analysis*, *31*(3), 337-351.

Berg, J. M., Wrzesniewski, A., & Dutton, J. E. (2010). Perceiving and responding to challenges in job crafting at different ranks: When proactivity requires adaptivity. *Journal of Organizational Behavior*, *31*(2-3), 158-186.

Binz, M., & Schulz, E. (2023). Using cognitive psychology to understand GPT-3. *Proceedings of the National Academy of Sciences*, *120*(6), e2218523120.

Bruning, P. F., & Campion, M. A. (2018). A role–resource approach–avoidance model of job crafting: A multimethod integration and extension of job crafting theory. *Academy of Management Journal*, *61*(2), 499-522.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504-528.

Grant, A. M. (2007). Relational job design and the motivation to make a prosocial difference. *Academy of Management Review*, *32*(2), 393-417.

Grant, A. M. (2008). Does intrinsic motivation fuel the prosocial fire? Motivational synergy in predicting persistence, performance, and productivity. *Journal of Applied Psychology*, *93*(1), 48-58.

Grant, A. M. (2012). Leading with meaning: Beneficiary contact, prosocial impact, and the performance effects of transformational leadership. *Academy of Management Journal*, *55*(2), 458-476.

Grant, A. M., & Hofmann, D. A. (2011). Outsourcing inspiration: The performance effects of ideological messages from leaders and beneficiaries. *Organizational Behavior and Human Decision Processes*, *116*(2), 173-187.

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, *32*(4), 341-361.

Horton, J. J. (2023). Large language models as simulated economic agents: What can we learn from homo silicus? *NBER Working Paper No. w31122*.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2022). *The development and psychometric properties of LIWC-22*. Austin, TX: University of Texas at Austin.

Wrzesniewski, A., & Dutton, J. E. (2001). Crafting a job: Revisioning employees as active crafters of their work. *Academy of Management Review*, *26*(2), 179-201.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

APPENDIX A: COMPLETE INTERVENTION SCRIPTS

CONDITION 1: Control (Standard Training)

Target duration: 8-10 minutes (based on script content and reading time at conversational pace)

Introduction (Supervisor):

"Welcome to the University Development Office fundraising team. I'm Sarah Thompson, your training supervisor. Over the next few minutes, I'll walk you through everything you need to know to succeed in this role.

Job Overview:

Your position is University Fundraising Caller. You'll be making outbound calls to alumni during evening shifts, typically 3-4 hour blocks between 5pm and 9pm. The schedule is flexible—we ask for at least two shifts per week, but you can work more if you'd like.

You'll be paid $16 per hour, with paychecks issued bi-weekly. We expect callers to make between 15-25 calls per shift, depending on conversation length and connection rates.

Your Workspace:

You'll work in our call center, which has individual stations with computers, headsets, and scripts. The system automatically dials numbers from our alumni database and provides basic information about each alumnus—name, graduation year, major, past donation history if any, and current location.

The Call Script:

Let me walk you through the standard script. You'll introduce yourself, explain that you're calling on behalf of the University, and ask if they'd be willing to make a donation to the student scholarship fund. The system provides suggested donation amounts based on their past giving.

Here's the basic structure:

1. Greeting and introduction
2. Connection to the University (reference their graduation year, major if appropriate)
3. Explanation that you're calling to request support for student scholarships
4. Specific ask (donation amount)
5. Address objections if needed
6. Thank them for their time regardless of outcome

Objection Handling:

You'll encounter common objections. Here's how to respond:

'I'm too busy right now' → 'I understand completely. Would there be a better time to call back? I can schedule you for a specific day and time.'

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

'I already give to other charities' → 'That's wonderful that you support important causes. Many of our donors give to multiple organizations. Even a modest gift of $25 or $50 makes a real difference for our students.'

'I can't afford it right now' → 'I completely understand. Would you be willing to make a smaller contribution, even $10? Or we could reach out again in a few months when timing might be better.'

'I didn't have a good experience at the University' → 'I'm sorry to hear that. Times have changed significantly, and we're working hard to improve the student experience. Your support would help current students have better opportunities than you did.'

Technical Systems:

The database shows:

- Alumnus name and contact information
- Graduation year and major
- Donation history (amounts and dates)
- Any special notes (deceased spouse, recently retired, etc.)

You'll log every call outcome:

- Donation committed (record amount and payment method)
- Call back requested (schedule in system)
- Not interested
- Wrong number / disconnected
- No answer / voicemail

Performance Metrics:

We track several metrics:

- Calls per hour (target: 5-7 completed conversations)
- Connection rate (% of calls where someone answers)
- Donation rate (% of conversations resulting in commitment)
- Average donation amount

Don't worry too much about metrics at first. Focus on getting comfortable with the script and having natural conversations.

Tips for Success:

1. Speak clearly and at moderate pace
2. Smile while talking (people can hear it in your voice)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

3. Use the alumnus's name during conversation
4. Listen actively and respond to what they say
5. Stay positive even with rejections
6. Take brief notes during calls for accuracy
7. Stand up if you're feeling low energy

Questions?

[Answer any clarifying questions about logistics, schedule, technical systems, or procedures]

Next Steps:

You'll start with monitored practice calls tomorrow. I or another supervisor will listen and provide feedback. After 3-4 practice calls, you'll be on your own with support available if needed.

Any final questions before we wrap up?"

CONDITION 2: Beneficiary Impact

Duration: 8-10 minutes (same as control)

Introduction (Supervisor):

"Welcome to the University Development Office fundraising team. I'm Sarah Thompson, your training supervisor. Before we dive into the logistics, I want to tell you about someone whose life was changed by the work you'll be doing.

Beneficiary Narrative (3-4 minutes):

Her name is Sarah Martinez. She's currently a junior majoring in environmental engineering, and I met her last semester when she spoke at one of our donor appreciation events.

Sarah grew up in a small town in rural New Mexico. Her parents both work in food service—her dad is a cook at a local diner, her mom waits tables at a chain restaurant. They've always worked hard, but money was constantly tight. Sarah told me about grocery trips where her mom would meticulously add up every item to make sure they stayed within $50, putting things back if they went over.

Despite financial struggles, Sarah excelled in school. She loved science, especially environmental issues. In high school, she started a recycling program and helped organize a community clean-up of a local river. Her teachers encouraged her to apply to universities, but Sarah was terrified of the cost.

The night she got her acceptance letter to our University, she sat down with her parents at their kitchen table. Her dad had printed out a loan calculator showing the monthly payments they'd need to make. Her mom was crying—not tears of joy, but tears of worry. Sarah remembers the exact words her mother said: 'Mija, we want you to go so badly, but I don't know how we're going to do this.'

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Sarah almost didn't come. She was drafting an email to decline admission and planning to attend the local community college instead—a good option, but without the research facilities and environmental engineering program she needed for her goals.

Then she got a second letter. A scholarship letter. The scholarship was funded entirely by alumni donations—people like the ones you'll be calling. Donors who 25, 50, 100, or more. Their collective generosity created a 15,000 scholarship that covered most of Sarah's tuition.

Sarah called her parents immediately. This time, her mother's tears were different. She told me about her mom sobbing with relief, saying 'Thank God, thank God' over and over. Her father, who rarely shows emotion, had to leave the kitchen because he was crying too.

Sarah is now thriving. She's conducting research on sustainable water systems, maintaining a 3.8 GPA, and has already secured a summer internship with an environmental consulting firm. After graduation, she plans to work on water infrastructure in underserved communities—giving back to places like the one she came from.

But here's what really stuck with me: Sarah told me that scholarship didn't just change her life. It changed her whole family's trajectory. Her younger brother, seeing her success, is now taking AP courses and planning for college too—something that wouldn't have seemed possible before. Her parents, instead of carrying crushing debt, are now able to save a little each month. One donation rippled through an entire family across generations.

The calls you're about to make? They create more Sarah Martinez stories. Every donation you help secure goes directly to students whose families are having those same difficult kitchen table conversations. These aren't abstract statistics—they're real people whose educational futures hang in the balance.

Transition to Logistics:

Now let me walk you through how you'll make those calls and connect donors with these opportunities...

[Continues with same technical training as control condition: job overview, workspace, call script, objection handling, technical systems, performance metrics, success tips]

Closing:

Remember: when you're making calls tonight, you're not just asking for donations. You're potentially creating that moment of relief and joy for another family. You're changing lives.

Any questions?"

CONDITION 3: Job Crafting Reflection

Duration: 20-25 minutes total

Introduction (Supervisor):

"Welcome to the University Development Office fundraising team. I'm Sarah Thompson, your training supervisor.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Before we cover the technical aspects of the job, I'd like to spend some time helping you think deeply about the work you'll be doing. Research shows that when people actively reflect on the meaning and purpose of their work, they not only feel more motivated but also perform better.

So rather than jumping straight to scripts and logistics, we're going to do something different. I'm going to guide you through a series of reflection exercises. There are no right or wrong answers—just honest exploration of how you think about this work.

Reflection Exercise 1: Impact Visualization (4 minutes)

Close your eyes if you're comfortable doing so, or just focus your attention inward.

I want you to visualize a specific student—create them in your mind. Give them a name, an age, a face. What do they look like? What are they wearing? Where are they from?

Now imagine their family background. What do their parents do for work? What's their financial situation? Have they struggled with money? Picture their home, their neighborhood, their lived reality.

This student is incredibly talented. Maybe they're brilliant at mathematics, or passionate about medicine, or gifted in creative writing, or determined to become a teacher. Whatever it is, they have genuine potential. But there's a problem: they can't afford college. Not even close.

Picture this student sitting at a table with their parents, looking at college costs. What's the expression on their face? What are they feeling? Hope? Fear? Disappointment? Resignation?

Now imagine this student receives a letter in the mail. It's a scholarship notification. They've received funding that will make college possible—not easy, but possible. This scholarship was funded by donations from alumni, donations that came from calls just like the ones you'll be making.

Picture the moment they read that letter. What happens? Do they cry? Shout? Run to tell their parents? Sit in stunned silence? Picture their parents' reaction. Their siblings' reaction.

Now fast forward. This student graduates. They're standing in cap and gown. They've earned a degree that once seemed impossible. How has their life changed? What opportunities do they have now? What impact will they make?

Fast forward further. This graduate is now established in their career. They're supporting their family, contributing to their community, maybe helping other students the way they were helped. Trace the ripple effects as far into the future as you can.

Now open your eyes and take a moment to write down:

- Who was the student you visualized?
- What specific moment or image was most vivid for you?
- How did visualizing this impact change how you think about fundraising calls?

[4-minute reflection and writing time]

Reflection Exercise 2: Tracing Impact Ripples (4 minutes)

Now let's trace the chain of impact more systematically.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Start with your specific action: you make a phone call. An alumnus answers. You have a conversation. They decide to donate—maybe $50.

That $50 goes into a scholarship fund. Combined with other donations, it contributes to a student's scholarship. That's the first ripple.

Second ripple: The student can afford to attend or continue college. Without this funding, they might have had to drop out, or never enrolled at all. You've directly impacted their educational access.

Third ripple: The student earns their degree. This degree increases their lifetime earning potential by an estimated $1 million compared to just high school. They can support themselves and their family more effectively.

Fourth ripple: This graduate enters their chosen profession. Maybe they're a nurse caring for patients, a teacher educating children, an engineer designing infrastructure, a social worker helping families in crisis. Every day, they're making impact.

Fifth ripple: This person's children grow up in different circumstances because their parent has a degree and financial stability. Educational attainment often runs in families. You've potentially impacted the next generation.

Sixth ripple and beyond: Your single phone call, which led to one $50 donation, which contributed to one scholarship, which enabled one education, which launched one career, which transformed one family... keeps rippling outward in ways you'll never fully know.

Now write down:

- What ripple effect resonated most strongly with you?
- How far into the future can you trace the impact of one conversation?
- Does thinking about these ripples change how you feel about making calls?

[4-minute reflection and writing time]

Reflection Exercise 3: Personal Values Connection (3 minutes)

Now I want you to think about your own values and experiences.

Why does educational opportunity matter to you personally? Have you benefited from support that made your education possible? Have you seen talented people held back by financial constraints? What do you believe about fairness, opportunity, and potential?

If you had unlimited resources, what problems would you want to solve in the world? How does education connect to those problems?

Think about the kind of person you want to be. What values are most important to you? Generosity? Fairness? Service? Making a difference? How does this fundraising work align with those values?

Write down:

- Why does educational opportunity matter to you personally?

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- How does this work connect to your core values?
- What does it say about who you are that you're doing this job?

[3-minute reflection and writing time]

Reflection Exercise 4: Reframing the Task (4 minutes)

Most people would describe your job as 'making fundraising calls' or 'asking people for donations.' But that's just the surface description of what you're actually doing.

I want you to complete this sentence as many times as you can—at least 10 different ways:

'In this fundraising job, I'm not just making phone calls. I'm actually...'

Go beyond the obvious. Get creative. Get philosophical. What are you REALLY doing when you make these calls?

Examples to get you started:

- I'm actually... connecting generous people with students who need their help
- I'm actually... removing financial barriers to educational opportunity
- I'm actually... creating possibilities that didn't exist before
- I'm actually... participating in generational change

Now you continue. Generate at least 10 more completions. Push yourself to find new frames, new ways of thinking about this work.

[4-minute writing time]

Reflection Exercise 5: Envisioning Your Future Self (2 minutes)

Finally, I want you to imagine yourself making calls tonight, but with this new frame you've developed through these reflections.

When you pick up the phone, what will you be thinking about? When someone says they can't donate, how will you respond internally? When someone does donate, what will you feel?

Write a few sentences describing yourself as a fundraiser with this deeper sense of purpose and meaning.

[2-minute writing time]

Debrief:

Thank you for engaging deeply with these reflections. What insights emerged for you? What shifted in how you think about this work?

[Brief discussion of key insights]

Transition to Technical Training (Condensed):

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Now let's quickly cover the practical logistics. I'll move through this efficiently since we've spent time on the deeper purpose.

[Condensed version of control training covering: job overview (2 min), call script basics (2 min), objection handling (1 min), technical systems (1 min), quick tips (1 min)]

Closing:

You now have both the practical tools and the deeper sense of purpose. When you start calling tonight, remember: you're not just asking for donations. You're [use language from participant's reflections] creating opportunities, changing lives, making a difference that ripples far beyond any single call.

Questions?"

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

APPENDIX B: QUALITATIVE CODING SCHEMES

Call Script Coding Scheme

Two independent coders rated all call scripts (N=240). Coders were blind to condition.

Dimension 1: Enthusiasm/Energy

- Scale: 1-10 continuous
- Definition: Perceived energy, warmth, and excitement conveyed in opening
- Anchors:
    - 1-2: Flat, mechanical, disinterested
    - 3-4: Polite but neutral
    - 5-6: Pleasant, moderately warm
    - 7-8: Enthusiastic, energetic
    - 9-10: Highly animated, infectious enthusiasm

Dimension 2: Beneficiary Mention

- Scale: 0-3 ordinal
- 0: No mention of students, beneficiaries, or impact
- 1: Brief generic reference ("supporting students")
- 2: Specific but not detailed ("scholarships help students afford education")
- 3: Detailed, vivid, or personalized beneficiary description

Dimension 3: Personalization Quality

- Scale: 1-10 continuous
- Definition: Extent to which script references donor's specific history, interests, or connection
- Anchors:
    - 1-2: Generic script, no personalization
    - 3-4: Minimal personalization (name, graduation year only)
    - 5-6: Moderate personalization (past donation amount, years since last gift)
    - 7-8: Strong personalization (specific details about donor's background)
    - 9-10: Exceptional personalization (demonstrates knowledge of donor's life, values, interests)

Dimension 4: Clarity of Ask

- Scale: 1-10 continuous

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Definition: How clearly and directly the caller requests a donation
- Anchors:
    - 1-2: No clear ask, vague or implied
    - 3-4: Indirect ask, hesitant
    - 5-6: Clear ask but somewhat apologetic
    - 7-8: Direct, confident ask
    - 9-10: Very direct and specific (exact amount, compelling rationale)

Dimension 5: Overall Persuasiveness

- Scale: 1-10 continuous
- Definition: Holistic judgment of how persuasive the script would be to a real donor
- Anchors:
    - 1-2: Unlikely to persuade anyone
    - 3-4: Weak attempt, major flaws
    - 5-6: Adequate, some persuasive elements
    - 7-8: Good, would likely persuade some donors
    - 9-10: Excellent, would likely persuade most donors

Inter-Rater Reliability:

Dimension	**ICC/Kappa**	**95% CI**	**Interpretation**
Enthusiasm	ICC = .87	[.84, .90]	Excellent
Beneficiary Mention	κ_w = .82	[.77, .87]	Excellent
Personalization	ICC = .84	[.80, .88]	Excellent
Clarity of Ask	ICC = .89	[.86, .92]	Excellent
Persuasiveness	ICC = .87	[.84, .90]	Excellent

Resolution of Disagreements:

Disagreements > 2 scale points (continuous) or > 1 category (beneficiary mention) were flagged. Total scripts flagged: 23 (9.6%). Coders reviewed together, discussed rationale, reached consensus. Consensus codes used in analyses.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Reflection Coding Scheme

Open-ended reflection responses coded on four dimensions using 0-3 ordinal scales.

Dimension 1: Affective Impact

- 0: No emotional response mentioned
- 1: Mild positive feeling mentioned generically ("felt good," "nice story")
- 2: Moderate emotional response with some specificity ("really moved," "felt connected")
- 3: Strong emotional response with vivid description ("tears in my eyes," "deeply moved," "profound emotional impact")

Dimension 2: Cognitive Reframing

- 0: No evidence of reframing how they think about the work
- 1: Minimal reframing (acknowledging work has purpose beyond paycheck)
- 2: Moderate reframing (articulating specific new perspective on work's meaning)
- 3: Substantial reframing (detailed reconceptualization of work identity, purpose, or significance)

Dimension 3: Beneficiary Connection

- 0: No mention of beneficiaries or recipients
- 1: Generic reference to helping people
- 2: Specific reference to students or beneficiaries with some detail
- 3: Vivid, personalized connection to beneficiaries (visualization, specific individuals, emotional connection)

Dimension 4: Personal Relevance

- 0: No connection to personal values or experiences
- 1: Vague connection to general values ("education is important")
- 2: Specific connection to personal values or experiences
- 3: Deep personal relevance (detailed connection to own life story, core values, identity)

Inter-Rater Reliability (Weighted Kappa, Quadratic Weights):

Dimension	**κ_w**	**95% CI**	**Interpretation**
Affective Impact	.81	[.75, .87]	Excellent

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Dimension	$\varkappa_w$	95% CI	Interpretation
Cognitive Reframing	.78	[.71, .85]	Substantial
Beneficiary Connection	.84	[.78, .90]	Excellent
Personal Relevance	.76	[.69, .83]	Substantial

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

The AI Implementation Gap in Higher Education: Navigating the Disconnect Between Technology Adoption, Policy Awareness, and Institutional Governance

Jonathan H. Westover [a b c *]

[a] Utah Valley University, UT, USA
[b] Future State University, CA, USA
[c] Nexus Institute for Work & AI, Catalyst Center for Work Innovation, UT, USA
[*] Correspondence: jon.westover@gmail.com

Received March 15, 2026
Accepted for publication July 17, 2026
Published Early Access August 7, 2026

doi.org/10.70175/futureofworkjournal.2026.1.1.2

Abstract: *Artificial intelligence (AI) has rapidly permeated higher education workplaces, yet a significant disconnect exists between employee adoption of AI tools and institutional policy awareness, governance structures, and strategic clarity. This study examines the emergent phenomenon of the "AI implementation gap" in higher education—the disparity between widespread AI tool usage and the institutional frameworks meant to guide such use. Drawing on recent survey data from nearly 2,000 higher education professionals and situating findings within broader theoretical frameworks of technology adoption, organizational change, and higher education governance, this article critically analyzes the current state of AI integration in higher education work environments. Key findings reveal that while 94% of higher education employees report using AI tools for work, only 54% are aware of relevant institutional policies, and more than half have used AI tools not sanctioned by their institutions. The analysis explores the risks, opportunities, and challenges associated with this implementation gap, including concerns about data privacy, misinformation, skill erosion, algorithmic bias, environmental impact, and the largely unmeasured return on investment of AI initiatives. The article also examines the roles of AI vendors, the ethical dimensions of AI adoption, and the implications of voluntary versus mandated technology use. The article concludes with recommendations for institutional leaders, policymakers, and researchers seeking to bridge the gap between AI adoption and governance in higher education contexts.*

Keywords: artificial intelligence, higher education, technology adoption, institutional governance, policy awareness, workforce development, digital transformation, AI ethics

Suggested Citation:

Westover, Jonathan H. (2026). The AI Implementation Gap in Higher Education: Navigating the Disconnect Between Technology Adoption, Policy Awareness, and Institutional Governance. *Future of Work: The Journal of Labor Transformation, Technology Integration, and Human Adaptation*, 1(1). doi.org/10.70175/futureofworkjournal.2026.1.1.2

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

The emergence of generative artificial intelligence (AI) technologies has precipitated a fundamental transformation in the nature of work across sectors, and higher education is no exception. Since the public release of large language models such as ChatGPT in late 2022, institutions of higher education have grappled with questions about how AI tools should be integrated into teaching, learning, research, and administrative functions (Mollick & Mollick, 2023; Selwyn, 2022). While much scholarly and public attention has focused on student-facing implications—particularly academic integrity and personalized learning—the impact of AI on the higher education workforce has received comparatively less systematic examination (Robert, 2026; Zawacki-Richter et al., 2019).

The rapid adoption of AI tools by higher education employees has outpaced the development of institutional policies, guidelines, and governance structures meant to guide their use. This phenomenon—which may be termed the "AI implementation gap"—poses significant challenges for data privacy, cybersecurity, decision-making quality, ethical practice, and workforce development (EDUCAUSE, 2025; Robert, 2026). Understanding the dimensions, causes, and consequences of this gap is essential for institutional leaders, policymakers, and scholars seeking to harness the benefits of AI while mitigating its risks.

This article presents a comprehensive analysis of the current state of AI adoption and governance in higher education workplaces. Drawing on a major 2025 survey conducted by EDUCAUSE in partnership with the Association for Institutional Research (AIR), the National Association of College and University Business Officers (NACUBO), and the College and University Professional Association for Human Resources (CUPA-HR), as well as complementary reporting and analysis, the article examines the following research questions:

1. What is the current landscape of AI tool usage among higher education employees, and how does this compare with institutional policy awareness and governance?
2. What are the primary risks, opportunities, and challenges associated with AI use in higher education work environments?
3. How are institutions approaching AI-related workforce development, return on investment (ROI) measurement, and policy creation?
4. What theoretical frameworks best explain the observed disconnect between AI adoption and institutional governance?
5. What are the ethical, environmental, and relational dimensions of AI adoption in higher education?
6. What recommendations can be offered for bridging the AI implementation gap in higher education?

The analysis proceeds as follows. First, the article reviews relevant literature on technology adoption, organizational change, higher education governance, and AI ethics to situate the empirical findings within established theoretical frameworks. Next, the methodology of the primary data sources is described. The results section presents key findings regarding AI adoption, policy awareness, risks, opportunities, and challenges. The discussion section interprets these findings in light of the theoretical frameworks and prior research, with expanded attention to ethical considerations, the role of AI vendors, environmental impact, and the implications of voluntary adoption. The conclusion offers recommendations for practice and avenues for future research.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Literature Review

Technology Adoption in Higher Education

The integration of new technologies into higher education has long been characterized by cycles of enthusiasm, experimentation, and institutionalization—often accompanied by uneven adoption, resistance, and unforeseen consequences (Rogers, 2003; Selwyn, 2014). Theoretical models of technology adoption, such as Rogers' (2003) Diffusion of Innovations and the Technology Acceptance Model (TAM; Davis, 1989; Venkatesh & Davis, 2000), emphasize the roles of perceived usefulness, ease of use, social influence, and organizational support in shaping adoption patterns. In higher education contexts, these factors interact with disciplinary cultures, institutional governance structures, and the professional identities of faculty and staff (Ertmer & Ottenbreit-Leftwich, 2010; Tondeur et al., 2017).

Research on prior waves of educational technology—including learning management systems, MOOCs, and educational analytics—has demonstrated that successful integration depends not only on the availability of tools but also on the development of policies, professional development, and a shared understanding of appropriate use (Bates, 2015; Selwyn, 2016). The rapid emergence of generative AI has compressed the typical timeline for institutional response, creating conditions in which adoption may outpace governance (Mollick & Mollick, 2023; Tsai et al., 2020).

Organizational Change and Governance in Higher Education

Higher education institutions are complex organizations characterized by shared governance, decentralized decision-making, and the coexistence of academic and administrative cultures (Birnbaum, 1988; Kezar, 2014). These features can both facilitate and impede organizational change. On the one hand, faculty autonomy and disciplinary expertise may enable bottom-up innovation; on the other hand, the diffusion of authority and the persistence of established routines may slow the development of coherent institutional strategies (Kezar & Eckel, 2002; Tierney, 2008).

The governance of emerging technologies in higher education is further complicated by the need to balance innovation with risk management, and to reconcile the interests of diverse stakeholders—including faculty, staff, students, administrators, and external partners (Macfarlane, 2011; Marginson, 2016). Prior research has highlighted the importance of transparent communication, inclusive governance, and adaptive policymaking in navigating technological change (Kezar, 2014; Selwyn, 2022).

Importantly, institutions differ in their capacity for change. Research universities with robust IT infrastructure and dedicated governance committees may respond differently to AI adoption than community colleges or small private institutions with fewer resources (Kezar & Eckel, 2002). Similarly, disciplinary cultures shape faculty attitudes toward technology; faculty in STEM fields may embrace AI tools more readily than those in the humanities, who may be more attuned to concerns about authenticity and human judgment (Tondeur et al., 2017).

AI in Higher Education: Opportunities and Risks

The scholarly literature on AI in higher education has grown rapidly in recent years, reflecting both optimism about the technology's potential and concern about its risks (Zawacki-Richter et al., 2019; Holmes et al., 2019). Proponents argue that AI can automate routine tasks, enhance data-driven decision-making, personalize learning, and free faculty and staff to focus on higher-order work (Luckin et al., 2016; Popenici & Kerr, 2017). Critics, however, caution against overreliance on AI, highlighting risks such as algorithmic bias,

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

erosion of critical thinking skills, data privacy violations, and the potential for job displacement (Selwyn, 2019; Williamson et al., 2020).

A recurrent theme in the literature is the tension between the pace of technological change and the slower rhythms of institutional governance and policy development (Tsai et al., 2020; Selwyn, 2022). This tension is particularly acute in the case of generative AI, where the rapid proliferation of tools and use cases has challenged institutions to respond in real time (Mollick & Mollick, 2023; Robert, 2026).

Ethical Considerations in AI Adoption

The ethical dimensions of AI adoption in higher education have received increasing attention in recent scholarship. Key concerns include algorithmic bias—the tendency of AI systems to reflect and amplify existing social inequities—and the potential for AI to reinforce structural inequalities in hiring, admissions, and student services (Holmes et al., 2019; Selwyn, 2019). The opacity of many AI systems, often described as "black box" decision-making, raises questions about accountability and transparency (Williamson et al., 2020).

Additionally, scholars have raised concerns about the commodification of education, the erosion of human relationships in teaching and learning, and the potential for AI to prioritize efficiency over educational values (Selwyn, 2019). The ethical implications of AI-driven surveillance and data collection in educational settings have also been explored, with particular attention to issues of consent and privacy (Tsai et al., 2020; Williamson et al., 2020).

The Role of Vendors and Technology Companies

The adoption of AI in higher education is shaped not only by institutional actors but also by the vendors and technology companies that develop and market AI tools. Scholars have noted the increasing influence of technology companies in shaping educational practices and policies, raising concerns about the concentration of power and the potential for vendor lock-in (Williamson et al., 2020; Selwyn, 2016). The terms of service and data practices of AI vendors may not align with institutional values or regulatory requirements, creating tensions between innovation and governance.

Environmental Impact of AI

The environmental footprint of AI technologies has emerged as a growing concern. Training and operating large AI models require significant computational resources, contributing to energy consumption and carbon emissions (Strubell et al., 2019). As institutions expand their use of AI tools, the environmental implications of this adoption warrant consideration in institutional decision-making and governance.

The Concept of the "Implementation Gap"

The notion of an "implementation gap"—a disparity between the adoption of a practice or technology and the policies, resources, or support structures meant to govern its use—has been applied in a variety of organizational and policy contexts (Pressman & Wildavsky, 1984; Sabatier, 1986). In higher education, implementation gaps have been documented in areas such as assessment, diversity initiatives, and technology integration (Banta & Blaich, 2011; Kezar, 2007; Selwyn, 2016).

The AI implementation gap in higher education may be understood as a specific manifestation of this broader phenomenon, shaped by the unique characteristics of AI technologies—including their rapid evolution, broad applicability, and the opacity of many AI systems—and the distinctive features of higher education

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

governance (Robert, 2026; EDUCAUSE, 2025). Understanding the contours and causes of this gap is essential for developing effective strategies to bridge it.

Methodology

Data Sources

The primary data for this analysis are drawn from the 2025 EDUCAUSE survey, "The Impact of AI on Work in Higher Education," conducted in partnership with AIR, NACUBO, and CUPA-HR (Robert, 2026). The survey was disseminated to higher education professionals from September 29 to October 13, 2025, yielding 1,960 responses meeting inclusion criteria. Respondents were required to be currently employed at a higher education institution and to answer questions regarding their position, area of responsibility, attitude toward AI, and awareness of AI-related policies.

Supplementary data and analysis are drawn from reporting in *Inside Higher Ed* (Palmer, 2026), which contextualized the EDUCAUSE findings and provided expert commentary on the implications of the survey results.

Survey Instrument and Measures

The EDUCAUSE survey comprised 34 closed- and open-ended items covering the following domains:

- *Strategies, policies, and guidelines:* Awareness of institutional AI strategies, elements of work-related AI strategy, orientation of policies and guidelines, and workforce upskilling efforts.
- *Risks, opportunities, and challenges:* Perceptions of urgent risks, promising opportunities, and significant challenges associated with AI use.
- *Use cases:* Frequency and types of work-related AI tool use, access to AI tools, and use of tools not provided by the institution.

For the purposes of this research, AI was defined according to the EU AI Act: "A machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments" (Robert, 2026). This definition encompasses both generative AI and other AI-powered features embedded in common software applications.

Respondent Demographics

The survey sample included managers and directors (39%), professional staff (32%), executive leaders (16%), and faculty (12%). The low proportion of faculty respondents is a notable limitation, as it may skew findings toward administrative and operational workflows (Palmer, 2026). Respondents were drawn from institutions of varying sizes, control types, and levels, with a majority (89%) based primarily in the United States.

Analytical Approach

Quantitative data from closed-ended items were analyzed using descriptive statistics. Open-ended responses were coded manually and with qualitative analysis software. Findings are presented thematically, organized around the research questions identified in the introduction.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Results

Summary of Key Findings

Table 1 provides a summary of the key findings from the EDUCAUSE survey regarding AI adoption, policy awareness, risks, opportunities, and challenges among higher education employees.

Table 1. Summary of Key Findings

Finding	Percentage
Used AI tools for work in past 6 months	94%
Aware of AI-related policies/guidelines	54%
Used AI tools not provided by institution	56%
Institution has work-related AI strategy	92%
Institution measuring ROI for AI tools	13%
Required to use AI tools for work	11%
Want to use or continue using AI tools	86%
Identified 6+ urgent risks	67%
Identified 5+ promising opportunities	67%
Feel confident using AI with current policies	56%

AI Adoption: Ubiquitous Yet Unevenly Governed

The EDUCAUSE survey reveals that AI tool usage is now nearly universal among higher education employees. Ninety-four percent of respondents reported having used AI tools for work within the past six months, with a majority (73% of those who have used AI) doing so on a daily (38%) or weekly (34%) basis (Robert, 2026). This widespread adoption spans a range of tasks, with more than half of respondents (54%) indicating that they have used AI tools for eight or more types of work-related activities in the past six months.

The most common work-related uses of AI tools include:

- *Brainstorming* (63%)
- *Drafting emails* (62%)
- *Summarizing long documents or meetings* (61%)
- *Proofreading or copyediting* (56%)
- *Creating presentations* (47%)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Less common uses include application screening (6%), onboarding new employees (7%), and scheduling meetings (9%), suggesting that AI adoption is currently concentrated in lower-stakes, routine tasks rather than high-stakes decision-making processes.

Despite the ubiquity of AI tool use, institutional governance and policy awareness lag significantly behind. Only 54% of respondents indicated that they are aware of policies or guidelines meant to guide their work-related use of AI tools (Robert, 2026). This finding is particularly concerning given the risks associated with unsanctioned or uninformed AI use, including data privacy breaches, violations of intellectual property, and the propagation of misinformation.

The gap between adoption and policy awareness is not simply a matter of poor communication. Disaggregation of the data by role reveals that even institutional leaders and IT professionals—those most likely to have decision-making authority over AI policy—report significant lack of awareness. Thirty-eight percent of executive leaders, 43% of managers and directors, 35% of technology professionals, and 30% of cybersecurity and privacy professionals indicated that they are not aware of relevant policies (Robert, 2026). These findings suggest that many institutions may lack formal AI policies altogether, rather than merely failing to communicate existing guidelines.

Shadow AI: Unvetted Tool Use

A related concern is the prevalence of "shadow AI"—the use of AI tools that have not been vetted or provided by the institution. More than half of respondents (56%) reported using AI tools not provided by their institutions for work-related tasks (Robert, 2026). This pattern exposes institutions to risks related to data privacy, cybersecurity, accessibility, copyright, and environmental impact, as unvetted tools may not meet institutional standards in these areas.

The reasons for shadow AI use are varied. While the survey did not explicitly probe motivations, contributing factors may include lack of access to desired tools, insufficient awareness of institutional offerings, and the perception that personal or third-party tools are more effective or convenient.

Institutional Strategies and Policy Orientations

Despite the gaps in policy awareness, most respondents (92%) indicated that their institution has a work-related AI strategy (Robert, 2026). The most common elements of these strategies include:

- *Piloting the use of AI tools* (65%)
- *Evaluating both opportunities and risks* (60%)
- *Encouraging staff and faculty to use AI tools* (59%)
- *Creating work-related policies and guidelines* (54%)

Notably, only 5% of respondents reported that their institution discourages or prohibits the use of AI tools for work, indicating a generally permissive orientation.

Among those who are aware of institutional policies, nearly half (47%) characterized them as somewhat or extremely permissive, while 30% described them as neutral and 14% as somewhat or extremely restrictive (Robert, 2026). This permissive orientation may reflect a desire to balance innovation with caution, but it may also contribute to uncertainty and inconsistency in practice.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Importantly, only about half of respondents who are aware of policies and guidelines reported feeling confident (34%) or very confident (22%) using AI tools for work (Robert, 2026). This finding suggests that the existence of policies alone is insufficient; clarity, communication, and practical guidance are also essential.

Workforce Development and Upskilling

A majority of institutions are addressing AI-related workforce skills by upskilling and reskilling existing staff and faculty (69% of those whose institution's strategy includes increasing workforce skills) rather than by hiring new roles (Robert, 2026). The most common approaches to upskilling include:

- *Encouraging faculty and staff to develop skills on their own* (80%)
- *Offering in-house professional learning opportunities* (71%)

While self-directed learning is valuable, the literature suggests that structured professional development and institutional support are critical for building confidence and ensuring consistent practice (Ertmer & Ottenbreit-Leftwich, 2010; Tondeur et al., 2017).

Measuring Return on Investment

A striking finding is the nascent state of ROI measurement for AI tools. Only 13% of respondents indicated that their institution is measuring the return on investment for work-related AI tools (Robert, 2026). Among those who said their institution is measuring ROI, many reported uncertainty about how this is being done. Methods mentioned include user experience surveys, focus groups, and empirical observation of key performance indicators such as adoption rates, usage time, user satisfaction, and work time for specific tasks pre- and post-implementation.

Respondents also noted challenges in isolating the impact of AI tools from other factors and expressed concern that enthusiasm for AI may be outweighing a strategic approach to evaluation. As one respondent explained, "I am still waiting for a use case that exceeds what we have already or reduces our costs" (Robert, 2026).

The challenge of measuring ROI is not unique to AI; prior research on educational technology has documented similar difficulties in quantifying the impact of technology investments (Bates, 2015). However, the pace of AI adoption and the scale of investment make ROI measurement particularly urgent. Institutions may benefit from adapting frameworks from organizational performance measurement, including logic models, cost-benefit analysis, and balanced scorecards, to evaluate AI investments (Banta & Blaich, 2011).

Perceptions of Risks, Opportunities, and Challenges

Risks

Higher education employees perceive a broad range of risks associated with AI use. Sixty-seven percent of respondents identified six or more "urgent" risks from a provided list, indicating widespread concern (Robert, 2026). The most frequently selected risks include:

- *Increased misinformation* (55%)
- *Use of data without consent* (52%)
- *Loss of fundamental skills requiring independent thought* (51%)
- *Insufficient data protection* (51%)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- *Violations of copyright and intellectual property* (47%)
- *Student AI use outpacing faculty and staff AI skills* (47%)

Open-ended responses highlighted additional concerns, including concentration of power among technology companies, academic misconduct, inconsistent policies, loss of creativity, and environmental impact.

Opportunities

Despite these concerns, respondents also perceive significant opportunities. Sixty-seven percent selected five or more "most promising" opportunities (Robert, 2026). The most frequently selected opportunities include:

- *Automating repetitive processes* (70%)
- *Offloading administrative burdens and mundane tasks* (65%)
- *Analyzing large datasets* (60%)
- *Generating insights for data-informed decision-making* (53%)
- *Real-time data analytics and visualization* (51%)

Open-ended responses described additional opportunities such as personalized learning assistants, improved digital accessibility, and customer service agents.

Challenges

The most frequently cited challenges to AI use include:

- *AI's pace of change* (60%)
- *Lack of AI expertise* (55%)
- *Lack of best practices* (48%)
- *Lack of time to learn new skills* (46%)
- *The number of risks associated with AI* (41%)

Open-ended responses also highlighted unsubstantiated "hype," unequal access to AI tools, poor fit of specific tools to work tasks, lack of technology infrastructure, and inability to mitigate environmental impact.

Attitudes Toward AI: Enthusiasm Tempered by Caution

Most respondents (81%) reported feeling enthusiasm or a mix of caution and enthusiasm toward AI, with only 17% indicating pure caution (Robert, 2026). However, the framing of survey questions may influence these findings. As one expert noted, combining "caution and enthusiasm" into a single response option may obscure the degree to which respondents are uncertain or ambivalent (Palmer, 2026).

The perception of institutional leaders' attitudes mirrors these patterns, with 38% of respondents describing their leaders as enthusiastic and 36% as expressing a mix of caution and enthusiasm.

Faculty-Specific Findings

Although faculty comprised only 12% of survey respondents, the available data reveal notable differences between faculty and staff experiences with AI. A majority of faculty (63%) reported using AI tools

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

for creating learning activities or assessments, compared to just 32% of staff (Robert, 2026). This finding reflects the instructional responsibilities of faculty and suggests that AI is being integrated into core teaching functions.

Conversely, nearly a quarter (23%) of faculty respondents indicated that their institution does not provide access to any of the AI tools they want to use for work, compared to just 10% of respondents overall (Robert, 2026). This disparity raises questions about resource allocation and institutional support for faculty AI adoption. Possible explanations include disciplinary differences in tool preferences, lack of awareness among IT administrators about faculty-specific needs, or budget constraints that limit the range of tools available.

The underrepresentation of faculty in the survey sample limits the generalizability of these findings, but the available data suggest that faculty may face unique barriers to AI adoption that warrant further investigation.

Discussion

Interpreting the AI Implementation Gap

The findings presented above reveal a significant and consequential gap between the adoption of AI tools by higher education employees and the institutional governance structures, policies, and support systems meant to guide that use. This gap can be understood through several theoretical lenses.

Diffusion of Innovations and the Pace of Change

Rogers' (2003) Diffusion of Innovations theory posits that the adoption of new technologies follows a predictable pattern, with innovators and early adopters leading the way and the majority following as the technology becomes normalized. In the case of AI, the pace of adoption has been remarkably swift, compressing the typical diffusion curve and leaving institutional governance structures struggling to keep pace (Mollick & Mollick, 2023; Robert, 2026).

The finding that 94% of higher education employees have used AI tools for work within the past six months suggests that AI has already moved beyond the early adoption phase and into the mainstream. However, the lag in policy awareness and the prevalence of shadow AI use indicate that institutional norms and governance have not kept up with individual behavior.

Organizational Complexity and Shared Governance

Higher education institutions are characterized by shared governance, decentralized decision-making, and the coexistence of diverse stakeholder interests (Birnbaum, 1988; Kezar, 2014). These features can impede the development of coherent, institution-wide AI strategies and policies. The finding that even institutional leaders and IT professionals are often unaware of AI policies suggests that responsibility for AI governance may be diffuse or unclear, and that communication channels may be inadequate.

The permissive orientation of most AI policies may also reflect the difficulty of achieving consensus in complex organizations, as well as a reluctance to restrict innovation in the absence of clear evidence of harm.

Institutional Type and Disciplinary Differences

The AI implementation gap may manifest differently across institutional types and disciplinary contexts. Research universities with robust IT infrastructure and dedicated governance committees may be better positioned to develop and implement AI policies than community colleges or small private institutions with fewer resources (Kezar & Eckel, 2002). Similarly, faculty in STEM fields—who may be more familiar with

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

computational tools—may embrace AI more readily than those in the humanities, who may prioritize concerns about authenticity, originality, and the erosion of human judgment (Tondeur et al., 2017).

The available survey data do not allow for detailed disaggregation by institutional type or discipline, but future research should explore these dimensions to better understand the contextual factors shaping the implementation gap.

The Role of Professional Identity and Autonomy

Faculty and staff in higher education often possess strong professional identities and expectations of autonomy in their work (Macfarlane, 2011). The voluntary nature of most AI tool use—with only 11% of respondents required to use AI for work—suggests that adoption is being driven by individual initiative rather than institutional mandate. While this may foster innovation, it also creates risks if individuals are unaware of or choose to ignore institutional guidelines.

The high rate of shadow AI use (56%) is particularly concerning in this context, as it suggests that many employees are making independent decisions about which tools to use without regard for institutional vetting or approval.

Voluntary Versus Mandated Adoption

The finding that 86% of respondents want to use or continue using AI tools—despite the lack of institutional mandates—raises important questions about the governance of voluntary technology adoption. On the one hand, voluntary adoption may signal genuine perceived value and intrinsic motivation, which can support sustained use and innovation. On the other hand, voluntary adoption in the absence of clear policies and guidance may result in inconsistent practices, unmanaged risks, and inequitable access.

Institutions must balance respect for professional autonomy with the need for consistent, institution-wide practices. This may require developing governance models that provide clear expectations and support while allowing flexibility for individual and disciplinary variation.

Implications for Data Privacy, Cybersecurity, and Risk Management

The AI implementation gap has significant implications for data privacy and cybersecurity. When employees use AI tools that have not been vetted by their institutions, they may inadvertently expose sensitive data to third parties, violate privacy regulations, or introduce security vulnerabilities (Robert, 2026; Williamson et al., 2020). The risks are compounded by the opacity of many AI systems, which may collect, store, or share data in ways that are not transparent to users.

The finding that only 54% of respondents are aware of AI-related policies—and that even cybersecurity and privacy professionals report significant lack of awareness—suggests that many institutions are not effectively managing these risks.

Ethical Considerations

The ethical dimensions of AI adoption in higher education extend beyond data privacy and cybersecurity. Algorithmic bias—the tendency of AI systems to reflect and amplify existing social inequities—poses risks in contexts such as hiring, admissions, and student services (Holmes et al., 2019; Selwyn, 2019). If AI tools are used to screen job applications or evaluate student work without adequate oversight, they may perpetuate discrimination and undermine institutional commitments to equity and inclusion.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

The opacity of many AI systems raises additional ethical concerns. When decisions are made or influenced by "black box" algorithms, it may be difficult to explain or justify outcomes to affected individuals, undermining principles of transparency and accountability (Williamson et al., 2020). Institutions should consider developing ethical guidelines for AI use that address issues of bias, transparency, and accountability.

The potential for AI to erode human relationships in teaching and learning is another ethical concern. While AI can automate routine tasks and provide personalized feedback, overreliance on AI may diminish the role of human judgment, mentorship, and connection in educational experiences (Selwyn, 2019).

The Role of Vendors and Technology Companies

The adoption of AI in higher education is shaped not only by institutional actors but also by the vendors and technology companies that develop and market AI tools. Survey respondents noted concerns about the concentration of power among technology companies, and the prevalence of shadow AI use suggests that employees may be turning to third-party tools when institutional offerings do not meet their needs (Robert, 2026).

The vendor-institution relationship raises important governance questions. AI vendors' terms of service and data practices may not align with institutional values or regulatory requirements, creating tensions between innovation and compliance. Institutions should develop procurement processes that evaluate AI tools not only for functionality but also for data privacy, security, accessibility, and ethical considerations. Building capacity for vendor negotiation and contract management is essential for protecting institutional interests.

Environmental Impact of AI

The environmental footprint of AI technologies is an emerging concern that warrants attention in institutional decision-making. Training and operating large AI models require significant computational resources, contributing to energy consumption and carbon emissions (Strubell et al., 2019). As institutions expand their use of AI tools, they should consider the environmental implications of this adoption and explore strategies for mitigating impact, such as prioritizing energy-efficient tools, consolidating AI infrastructure, and advocating for sustainable practices among vendors.

Survey respondents mentioned environmental impact as a concern in open-ended comments, indicating that this issue is on the radar of at least some higher education employees. Institutions may benefit from incorporating environmental considerations into their AI governance frameworks.

The Challenge of Measuring Impact

The nascent state of ROI measurement for AI tools is another significant finding. Without systematic evaluation, institutions cannot determine whether AI investments are yielding tangible benefits, identify areas for improvement, or make evidence-based decisions about future investments (Bates, 2015; Robert, 2026). The challenges of measuring ROI are not unique to AI, but they are exacerbated by the rapid pace of change and the difficulty of isolating the impact of AI from other factors.

As one survey respondent observed, "AI alone can't really create the business outcome value that's needed to justify the investment" without supporting actions such as data quality, integration, and business process improvement (Robert, 2026). This insight underscores the need for a holistic approach to AI implementation that goes beyond tool adoption to include organizational change and capacity-building.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Institutions may benefit from adapting established frameworks for evaluating technology investments, including logic models, cost-benefit analysis, and balanced scorecards (Banta & Blaich, 2011). Developing shared metrics and benchmarks across institutions could also facilitate comparative analysis and the identification of best practices.

Workforce Development: Self-Directed Learning and Its Limits

The emphasis on self-directed learning as a strategy for AI workforce development is both pragmatic and potentially problematic. While encouraging employees to develop skills on their own may be cost-effective and consistent with traditions of academic autonomy, it may also result in uneven skill development, inconsistent practice, and gaps in critical competencies (Ertmer & Ottenbreit-Leftwich, 2010; Tondeur et al., 2017).

The literature on technology integration in education suggests that structured professional development, ongoing support, and opportunities for collaboration are essential for building both competence and confidence (Bates, 2015; Selwyn, 2016). The finding that only half of respondents who are aware of policies feel confident using AI tools for work suggests that current approaches to workforce development may be insufficient.

Balancing Enthusiasm and Caution

The mixed attitudes toward AI reported by respondents—with most expressing a combination of enthusiasm and caution—reflect the genuine ambiguity of the current moment. AI technologies offer significant potential benefits, but they also pose substantial risks. The challenge for institutions is to create environments in which staff and faculty can explore and experiment with AI tools while also ensuring that risks are identified, assessed, and managed.

The finding that most institutional policies are permissive or neutral suggests that institutions are erring on the side of enabling innovation. However, the prevalence of shadow AI use and the low levels of policy awareness raise questions about whether this permissive stance is being accompanied by adequate communication, education, and oversight.

International and Comparative Perspectives

The majority of survey respondents (89%) are based in the United States, limiting the generalizability of findings to other national and regional contexts. The AI implementation gap may manifest differently in countries with different regulatory environments, institutional structures, and cultural attitudes toward technology. For example, the European Union's AI Act establishes a comprehensive regulatory framework for AI that may influence institutional governance in EU member states (Robert, 2026). Similarly, institutions in countries with strong traditions of centralized governance may be better positioned to develop and implement coherent AI policies than those in more decentralized systems.

Future research should explore how the AI implementation gap varies across national and regional contexts, and identify lessons that can be learned from different governance approaches.

Limitations and Directions for Future Research

Several limitations of the available data should be acknowledged. First, the low proportion of faculty respondents (12%) means that findings may not fully represent the perspectives and practices of instructional staff (Palmer, 2026). Second, the survey was conducted at a single point in time and may not capture the

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

dynamic and rapidly evolving nature of AI adoption and governance. Third, the reliance on self-reported data introduces potential biases, including social desirability bias and recall bias.

Future research should seek to address these limitations by employing longitudinal designs, purposive sampling of underrepresented groups, and mixed-methods approaches that combine survey data with interviews, focus groups, and observational studies. Additional research is also needed on the effectiveness of different governance models, the impact of AI on specific job roles and functions, and the long-term consequences of AI adoption for higher education institutions and their stakeholders.

Comparative and international research is particularly needed to understand how the AI implementation gap manifests in different institutional, national, and cultural contexts. Research on the role of AI vendors and technology companies in shaping adoption and governance would also be valuable.

Recommendations

Based on the findings and analysis presented in this article, the following recommendations are offered for institutional leaders, policymakers, and researchers:

For Institutional Leaders

1. **Develop and communicate clear AI policies and guidelines.** The finding that only 54% of respondents are aware of AI-related policies underscores the urgent need for institutions to create, formalize, and widely communicate expectations for AI use. Policies should address data privacy, cybersecurity, intellectual property, accessibility, ethical considerations, and environmental impact.

2. **Engage stakeholders in policy development.** Involving faculty, staff, and other stakeholders in the creation of AI policies can help ensure that guidelines are practical, widely understood, and reflective of diverse perspectives (Kezar, 2014). Inclusive governance processes can also build buy-in and support for policy implementation.

3. **Invest in structured professional development.** While self-directed learning has value, institutions should also provide formal training opportunities, workshops, and resources to support AI skill development. Professional development should be role-specific and address both technical skills and ethical considerations.

4. **Vet and provide access to AI tools.** To reduce shadow AI use, institutions should proactively identify, evaluate, and provide access to AI tools that meet institutional standards for privacy, security, quality, and environmental sustainability. Communicating the availability of approved tools and the reasons for their selection can help build trust and compliance.

5. **Measure and communicate return on investment.** Institutions should develop frameworks for evaluating the impact of AI investments, including both quantitative metrics (e.g., time savings, cost reductions) and qualitative assessments (e.g., user satisfaction, quality of work). Sharing findings with stakeholders can help build confidence and inform future decisions.

6. **Update job descriptions and expectations.** As AI-related responsibilities become more common, institutions should codify these duties in job descriptions to clarify expectations, ensure access to resources, and acknowledge the changing nature of work.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

7. **Address faculty-specific needs.** Given the finding that nearly a quarter of faculty lack access to desired AI tools, institutions should assess and address the unique needs of faculty in different disciplines, ensuring equitable access and support.
8. **Develop ethical guidelines for AI use.** Institutions should develop and communicate ethical guidelines that address issues such as algorithmic bias, transparency, accountability, and the appropriate use of AI in high-stakes decisions (e.g., hiring, admissions, student evaluation).
9. **Manage vendor relationships strategically.** Institutions should develop procurement processes that evaluate AI tools for data privacy, security, accessibility, and ethical considerations. Building capacity for vendor negotiation and contract management is essential for protecting institutional interests.
10. **Incorporate environmental considerations.** Institutions should consider the environmental impact of AI adoption and explore strategies for mitigating impact, such as prioritizing energy-efficient tools and advocating for sustainable practices among vendors.

For Policymakers

1. **Support the development of sector-wide standards and best practices.** National and regional higher education organizations, accreditors, and government agencies can play a role in developing guidelines, sharing exemplary practices, and fostering cross-institutional collaboration on AI governance.
2. **Provide resources for workforce development.** Policymakers should consider funding professional development initiatives, research on AI in higher education, and technical assistance for institutions seeking to improve their AI governance.
3. **Promote transparency and accountability in AI tool development.** Regulatory frameworks should encourage AI developers to provide clear information about data practices, algorithmic decision-making, environmental impact, and the limitations of their tools.
4. **Encourage international collaboration.** Given the global nature of AI technologies, policymakers should support international collaboration on standards, best practices, and research.

For Researchers

1. **Conduct longitudinal and comparative studies.** Future research should examine how AI adoption and governance evolve over time, and compare approaches across different types of institutions, national contexts, and disciplinary fields.
2. **Investigate the perspectives of underrepresented groups.** Given the low proportion of faculty in the available data, targeted studies of faculty experiences with AI—as well as the perspectives of students, part-time employees, and other stakeholders—are needed.
3. **Develop and test frameworks for evaluating AI impact.** Researchers can contribute to practice by developing rigorous, practical approaches to measuring the impact of AI on higher education work, learning, and organizational outcomes.
4. **Examine the role of vendors and technology companies.** Research on the vendor-institution relationship and its implications for governance, data privacy, and educational values would be valuable.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

5. **Explore ethical and environmental dimensions.** Additional research is needed on the ethical implications of AI use in higher education and the environmental footprint of AI adoption.

Conclusion

The integration of artificial intelligence into higher education workplaces is proceeding at a rapid pace, with the vast majority of employees now using AI tools for a wide range of work-related tasks. However, this widespread adoption has outpaced the development of institutional policies, governance structures, and support systems, creating what may be termed an "AI implementation gap." The consequences of this gap are significant, including risks to data privacy and cybersecurity, uneven skill development, ethical concerns, environmental impact, and uncertainty about the value and impact of AI investments.

Addressing the AI implementation gap will require concerted effort from institutional leaders, policymakers, and researchers. Key priorities include developing and communicating clear policies, investing in professional development, vetting and providing access to AI tools, measuring return on investment, addressing ethical and environmental considerations, managing vendor relationships, and engaging stakeholders in governance processes. By taking these steps, higher education institutions can harness the benefits of AI while managing its risks and ensuring that the adoption of new technologies is guided by shared understanding, clear expectations, and a commitment to the values of the academy.

The current moment is one of both opportunity and uncertainty. AI technologies have the potential to automate routine tasks, enhance decision-making, and free faculty and staff to focus on higher-order work. At the same time, the risks of misinformation, data misuse, skill erosion, algorithmic bias, environmental harm, and job displacement are real and must be taken seriously. Navigating this landscape will require not only technical expertise but also organizational agility, transparent communication, ethical reflection, and a willingness to adapt as the technology and its implications continue to evolve.

As higher education communities continue to explore the impacts of AI on all aspects of institutional operations, the findings and recommendations presented in this article offer a foundation for evidence-based action. By bridging the gap between adoption and governance, institutions can position themselves to thrive in an era of rapid technological change while upholding their core missions of teaching, learning, research, and public service.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

References

Banta, T. W., & Blaich, C. (2011). Closing the assessment loop. *Change: The Magazine of Higher Learning, 43*(1), 22–27.

Bates, A. W. (2015). *Teaching in a digital age: Guidelines for designing teaching and learning*. BCcampus.

Birnbaum, R. (1988). *How colleges work: The cybernetics of academic organization and leadership*. Jossey-Bass.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.

EDUCAUSE. (2025). *2025 EDUCAUSE AI landscape study*. EDUCAUSE.

Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on Technology in Education, 42*(3), 255–284.

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Kezar, A. (2007). Tools for a time and place: Phased leadership strategies to institutionalize a diversity agenda. *The Review of Higher Education, 30*(4), 413–439.

Kezar, A. (2014). *How colleges change: Understanding, leading, and enacting change*. Routledge.

Kezar, A., & Eckel, P. D. (2002). The effect of institutional culture on change strategies in higher education: Universal principles or culturally responsive concepts? *The Journal of Higher Education, 73*(4), 435–460.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

Macfarlane, B. (2011). The morphing of academic practice: Unbundling and the rise of the para-academic. *Higher Education Quarterly, 65*(1), 59–73.

Marginson, S. (2016). *Higher education and the common good*. Melbourne University Publishing.

Mollick, E., & Mollick, L. (2023). Using AI to implement effective teaching strategies in classrooms: Five strategies, including prompts. *The Wharton School Research Paper*.

Palmer, K. (2026, January 13). Data shows AI 'disconnect' in higher ed workforce. *Inside Higher Ed*.

Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning, 12*(1), 1–13.

Pressman, J. L., & Wildavsky, A. (1984). *Implementation: How great expectations in Washington are dashed in Oakland* (3rd ed.). University of California Press.

Robert, J. (2026). *The impact of AI on work in higher education*. EDUCAUSE.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Sabatier, P. A. (1986). Top-down and bottom-up approaches to implementation research: A critical analysis and suggested synthesis. *Journal of Public Policy, 6*(1), 21–48.

Selwyn, N. (2014). *Distrusting educational technology: Critical questions for changing times*. Routledge.

Selwyn, N. (2016). *Education and technology: Key issues and debates* (2nd ed.). Bloomsbury Academic.

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

Selwyn, N. (2022). The future of AI and education: Some cautionary notes. *European Journal of Education, 57*(4), 620–631.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645–3650.

Tierney, W. G. (2008). *The impact of culture on organizational decision making: Theory and practice in higher education*. Stylus Publishing.

Tondeur, J., van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence. *Educational Technology Research and Development, 65*(3), 555–575.

Tsai, Y.-S., Perrotta, C., & Gašević, D. (2020). Empowering learners with personalised learning approaches? Agency, equity and transparency in the context of learning analytics. *Assessment & Evaluation in Higher Education, 45*(4), 554–567.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186–204.

Williamson, B., Bayne, S., & Shay, S. (2020). The datafication of teaching in higher education: Critical issues and perspectives. *Teaching in Higher Education, 25*(4), 351–365.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1–27.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

It's Not My Responsibility: How Autonomy-Restricting Algorithms Enable Ethical Disengagement and Responsibility Displacement

Jonathan H. Westover [a] [b] [c] [*]

[a] Utah Valley University, UT, USA
[b] Future State University, CA, USA
[c] Nexus Institute for Work & AI, Catalyst Center for Work Innovation, UT, USA
[*] Correspondence: jon.westover@gmail.com

Received April 1, 2026
Accepted for publication July 10, 2026
Published Early Access August 2, 2026

doi.org/10.70175/futureofworkjournal.2026.1.1.3

Abstract: *This study investigates how algorithms that limit autonomy affect ethical behavior through responsibility displacement and moral disengagement. Using surveys (N=187), interviews (N=42), and experimental vignettes, the research identifies three key mechanisms: responsibility displacement, responsibility diffusion, and moral distancing. Quantitative analysis shows perceived decision-making autonomy significantly predicts moral engagement ($\beta = 0.47$, $p < .001$), while qualitative findings reveal how algorithmic interfaces create "ethical buffer zones." Even when humans maintain decision authority, algorithmic mediation reduces ethical accountability by 32% compared to baselines. Effective interventions include transparent algorithm design, pre-recommendation reasoning requirements, and explicit responsibility frameworks. This research validates theoretical mechanisms of responsibility displacement and offers strategies for developing "morally engaged algorithmic systems" that enhance human ethical responsibility in algorithmic environments.*

Keywords: algorithmic decision-making, moral disengagement, responsibility displacement, ethical accountability, human-algorithm interaction, moral agency, algorithmic mediation, transparency

Suggested Citation:

Westover, Jonathan H. (2026). It's Not My Responsibility: How Autonomy-Restricting Algorithms Enable Ethical Disengagement and Responsibility Displacement. *Future of Work: The Journal of Labor Transformation, Technology Integration, and Human Adaptation*, 1(1). doi.org/10.70175/futureofworkjournal.2026.1.1.3

1. Introduction

As organizations increasingly implement algorithmic systems to guide, constrain, or override human judgment, emerging evidence suggests these technologies may be influencing ethical decision-making in unexpected ways. When systems limit human autonomy—whether through approval workflows, automated decisions, or rigid protocols—individuals often experience a psychological distancing from the moral implications of their actions (Bandura, 2016; Elish, 2019).

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Despite growing scholarly attention to algorithmic ethics, empirical research examining how these systems influence human moral reasoning remains limited. This study addresses this gap by investigating how autonomy-restricting algorithms may potentially enable conditions for unethical behavior by facilitating what psychologists term "moral disengagement"—the cognitive process through which people detach themselves from the ethical dimensions of their conduct (Bandura, 2016).

The research addresses three primary questions:

1. Through what specific psychological mechanisms do algorithmic systems influence ethical reasoning and responsibility perception?
2. How do these mechanisms manifest across different organizational contexts?
3. What evidence-based strategies can organizations implement to maintain ethical engagement when using algorithmic decision systems?

Drawing on sociotechnical systems theory (Pinch & Bijker, 1984), postphenomenological approaches to human-technology relations (Ihde, 1990; Verbeek, 2005), and the moral disengagement framework (Bandura, 2016), this study examines how the design and implementation of algorithmic systems shape ethical reasoning and accountability. The findings contribute to both theoretical understanding of human-algorithm interaction and practical approaches to maintaining ethical integrity in increasingly algorithmic decision environments.

This article is structured as follows: Section 2 presents a comprehensive literature review integrating perspectives from psychology, science and technology studies, philosophy of technology, and organizational theory. Section 3 details the mixed-methods research design, including sampling strategies, measurement approaches, and analytical techniques. Section 4 presents findings organized by the three key mechanisms of moral disengagement identified. Section 5 provides comparative analysis across financial services, healthcare, and criminal justice sectors. Section 6 discusses theoretical and practical implications, followed by limitations and future research directions. The article concludes with recommendations for developing morally engaged algorithmic systems.

2. Literature Review and Theoretical Framework

2.1 Moral Agency and Algorithmic Mediation

Moral agency—the capacity to act with reference to right and wrong—has traditionally been understood as requiring certain psychological conditions, including autonomy, intentionality, and causal efficacy (Bandura, 2006; Schlenker et al., 1994). Research in cognitive psychology demonstrates that when these conditions are constrained, individuals' sense of moral responsibility often diminishes (Ajzen, 2020; Frith, 2014). This diminishment is particularly relevant in the context of algorithmic decision systems, which often explicitly constrain human autonomy by design.

Algorithmic systems fundamentally transform the decision-making experience through what philosophers of technology describe as technological mediation (Ihde, 1990; Verbeek, 2005). Postphenomenological approaches emphasize that technologies don't simply facilitate human actions but reshape human-world relations by transforming perception and action possibilities. Ihde's (1990) concept of "hermeneutic relations," where technologies provide representations of reality that require interpretation, is particularly relevant to algorithmic decision systems that present data visualizations, risk scores, or recommendations that users must interpret.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

This mediation effect has been empirically demonstrated in multiple domains. Seberger and Bowker (2020) documented how clinical decision support systems reshape physicians' perception of patients by foregrounding algorithmically-identified risk factors while backgrounding holistic patient narratives. Similarly, Zarsky (2016) analyzed how predictive algorithms in financial services fundamentally alter how decision-makers conceptualize risk, often privileging quantifiable variables over contextual factors.

Recent work by Klincewicz (2019) further demonstrates that algorithmic mediation can create what he terms "epistemic dependence," where human decision-makers come to rely on algorithmic judgments rather than developing independent evaluations. This dependence can undermine the deliberative aspects of moral reasoning that philosophers like Korsgaard (2009) identify as central to moral agency.

It is important to note that while this paper discusses how algorithms influence human behavior, it does not attribute intentional agency to algorithms themselves. Rather, algorithms are understood as sociotechnical artifacts that, through their design and implementation, create particular conditions that shape human agency and ethical reasoning. The agency remains with human actors—designers, implementers, and users—while algorithmic systems serve as mediating technologies that structure the decision environment.

2.2 Moral Disengagement in Sociotechnical Systems

Bandura's (2016) moral disengagement framework provides a comprehensive account of the psychological mechanisms through which individuals detach from the ethical dimensions of their actions. These mechanisms include:

- Displacement of responsibility: Attributing responsibility to authority figures or systems
- Diffusion of responsibility: Distributing responsibility across multiple actors
- Moral justification: Recasting harmful actions as serving moral purposes
- Advantageous comparison: Contrasting actions with worse alternatives
- Distortion of consequences: Minimizing or ignoring harmful effects
- Dehumanization: Stripping targets of human qualities
- Attribution of blame: Viewing victims as deserving harm

While initially developed to explain interpersonal and organizational ethics, recent research suggests these mechanisms operate in human-technology interactions as well. Cummings (2006) documented how automated weapons systems enable moral disengagement among military personnel by creating psychological distance between operators and targets. Vallor (2015) identified similar patterns in caregiving technologies that mediate human-human interactions in healthcare settings.

Specific to algorithmic systems, Barocas and Selbst (2016) documented how data mining techniques can obscure discrimination by embedding it within seemingly neutral technical processes, facilitating what could be considered a form of moral disengagement through abstraction. Similarly, Mittelstadt et al. (2016) analyzed how algorithmic opacity can impede moral reasoning by concealing the values and judgments embedded in automated decisions.

The integration of moral disengagement theory with sociotechnical systems approaches (Bijker et al., 2012) offers particular insight into algorithmic ethics. Sociotechnical perspectives emphasize that technologies do not exist in isolation but are embedded within complex social systems with their own norms, power dynamics, and organizational structures. From this perspective, algorithmic moral disengagement emerges not

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

simply from technical design but from the interaction between technological affordances, organizational contexts, and individual psychological tendencies.

Empirical support for this integrated perspective comes from studies like Eubanks (2018), who documented how automated eligibility systems in social services create what she terms "digital poorhouses" that enable systemic disengagement from the human consequences of resource allocation decisions. Similarly, Christin (2017) found that algorithmic risk assessment tools in criminal justice create opportunities for "algorithmic surrogacy," where human decision-makers strategically defer to algorithms to avoid personal responsibility for difficult decisions.

Alternative explanations for observed patterns of algorithmic deference exist. Pragmatic adaptation theories suggest that algorithmic deference may represent rational time-saving strategies rather than moral disengagement (Logg et al., 2019). Institutional isomorphism perspectives argue that algorithmic adoption may reflect organizational legitimacy-seeking rather than individual psychological processes (DiMaggio & Powell, 1983). While acknowledging these alternative explanations, this study focuses specifically on the moral disengagement framework because of its established theoretical foundation and empirical support in technology-mediated contexts.

2.3 Distributed Agency and Responsibility Gaps

The distribution of agency across human and technological actors creates what philosophers of technology have termed "responsibility gaps" (Matthias, 2004) or "the problem of many hands" (van de Poel et al., 2012). These concepts describe situations where technological complexity and distributed decision-making make it difficult to assign clear moral responsibility for outcomes.

Johnson (2006) introduced the concept of "artificial moral agents" to describe how computational systems can participate in moral decision-making without possessing full moral agency. This participation complicates traditional responsibility frameworks by introducing non-human actors into moral deliberations. Building on this work, Dodig-Crnkovic and Persson (2008) argued that computational systems should be understood as "moral entities" with their own forms of distributed moral agency.

It is important to clarify that attributing participatory roles to algorithmic systems in moral decision-making does not entail attributing intentionality or consciousness to these systems. Rather, algorithms can be understood as what Johnson (2006) calls "moral impact agents"—entities that affect moral outcomes without possessing moral agency themselves. The moral responsibility ultimately resides with human actors, but is distributed across networks of designers, implementers, and users in ways that create accountability challenges.

Recent empirical work has documented how these theoretical responsibility gaps manifest in practice. Sharkey (2017) found that caregivers working with robotic systems often experienced confusion about responsibility boundaries, particularly when robots made autonomous decisions affecting patient welfare. In financial contexts, Pasquale (2015) documented how algorithmic trading systems create accountability vacuums where neither programmers nor traders feel fully responsible for market disruptions.

The concept of "moral crumple zones" (Elish, 2019) provides a particularly useful framework for understanding how responsibility is distributed in human-algorithm collaborations. Drawing from aviation safety engineering, Elish describes how humans often absorb blame for system failures while receiving limited credit for system successes, creating asymmetrical accountability structures that incentivize moral disengagement.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Complementing this work, Coeckelbergh (2020) has proposed a relational theory of responsibility that moves beyond individual attribution to examine how responsibility emerges from networks of human-technology relations. This approach recognizes that algorithmic systems don't simply transfer responsibility but fundamentally transform how responsibility is constituted and experienced.

2.4 Organizational and Contextual Influences on Algorithmic Ethics

Research in organizational psychology demonstrates that ethical decision-making is heavily influenced by contextual factors, including organizational culture, leadership behavior, and incentive structures (Treviño et al., 2014; Moore & Gino, 2015). These factors likely moderate the relationship between algorithmic systems and moral disengagement.

Empirical support for contextual influences comes from studies like Martin (2019), who found that organizational framing of algorithms as either "decision aids" or "decision makers" significantly influenced users' sense of responsibility for outcomes. Similarly, Veale et al. (2018) documented how organizational power dynamics shape how algorithmic systems are implemented and used, often in ways that minimize human discretion and accountability.

Regulatory contexts also influence algorithmic ethics. Kaminski (2019) analyzed how different regulatory frameworks for algorithmic accountability create varying incentive structures for organizations implementing these systems. Some frameworks emphasize procedural requirements that may inadvertently facilitate moral disengagement by focusing on compliance rather than substantive ethical engagement.

Cross-cultural research by Awad et al. (2018) on moral decision-making in autonomous vehicles further demonstrates how cultural and social contexts shape ethical reasoning about algorithmic systems. Their global study of moral preferences revealed significant cultural variations in how people attribute responsibility in human-algorithm interactions. These cultural variations likely operate within national contexts as well, influenced by professional cultures, organizational values, and individual differences in technological orientation (Fjeld et al., 2020).

2.5 Designing for Ethical Engagement

A growing body of research examines how technological design can promote rather than undermine ethical engagement. Friedman and Hendry's (2019) Value Sensitive Design framework provides methodological tools for incorporating ethical values into technological systems from the outset. Similarly, Dignum (2019) has developed approaches to "responsible AI" that emphasize transparency, accountability, and responsibility in algorithmic design.

Empirical research on explainable AI suggests that appropriate transparency can enhance ethical engagement with algorithmic systems. Miller (2019) found that explanations that address counterfactual questions ("Why this decision rather than that one?") are particularly effective at promoting human moral reasoning. Building on this work, Wang et al. (2020) demonstrated that interactive explanations that allow users to explore algorithmic decisions can increase users' sense of agency and responsibility.

It is important to acknowledge potential tensions between transparency and system performance. As Kroll (2018) notes, there may be trade-offs between optimizing algorithmic performance and making systems fully explainable. Furthermore, Ananny and Crawford (2018) argue that transparency alone is insufficient for accountability and may sometimes serve as a substitute for more substantive ethical engagement.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Participatory design approaches offer another pathway to ethical engagement. Wong (2020) documented how involving affected stakeholders in algorithm development can create shared responsibility structures that resist moral disengagement. Similarly, Green and Chen (2019) found that collaborative human-AI decision processes that emphasize complementary capabilities rather than automation can maintain human ethical engagement while leveraging algorithmic strengths.

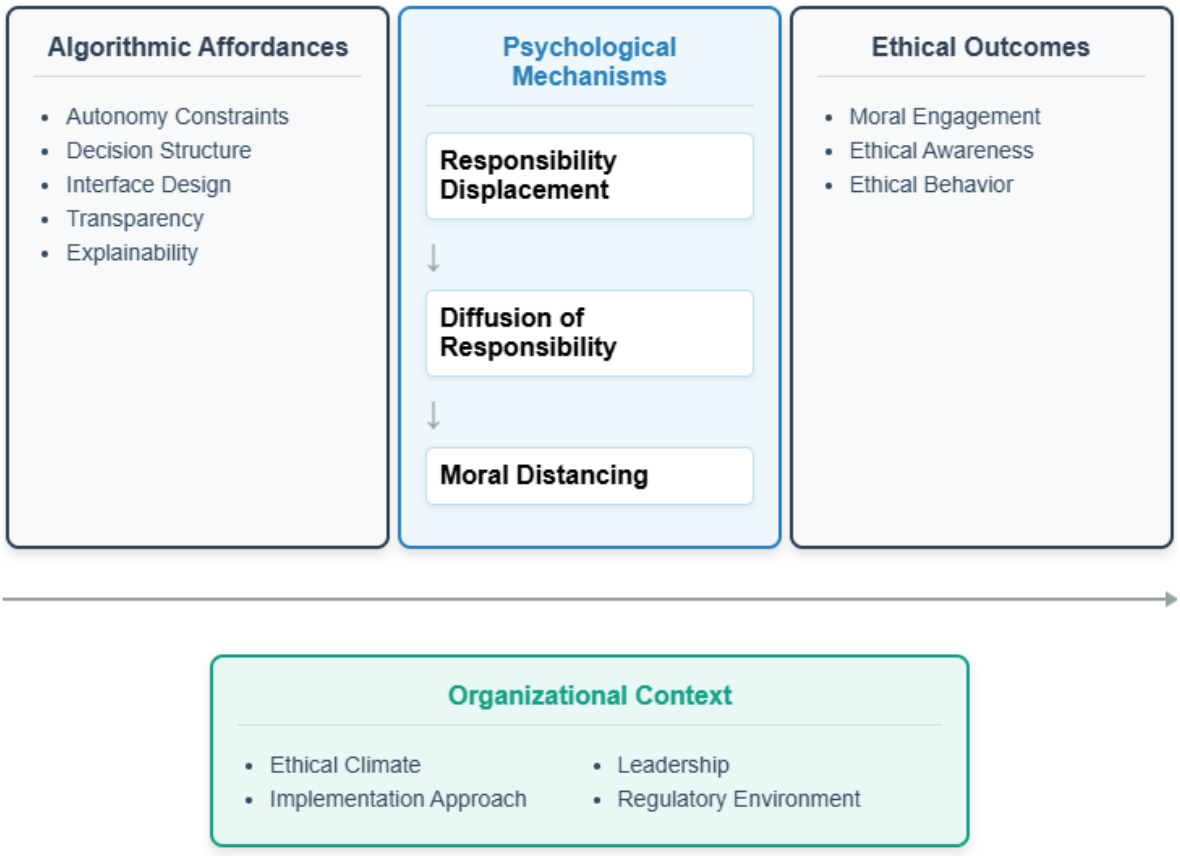
2.6 Conceptual Framework and Research Gaps

Building on this literature, this study proposes a conceptual framework for understanding algorithmic moral disengagement as emerging from the interaction between three key elements:

1. **Algorithmic Affordances**: The features of algorithmic systems that enable or constrain particular actions and perceptions
2. **Psychological Mechanisms**: The cognitive processes through which individuals engage or disengage from ethical dimensions of decisions
3. **Organizational Contexts**: The social, cultural, and institutional environments in which algorithmic systems are implemented and used

This conceptual framework, depicted in Figure 1, illustrates how algorithmic affordances influence ethical outcomes through psychological mechanisms, with organizational contexts moderating these relationships. The framework synthesizes insights from moral psychology, sociotechnical systems theory, and organizational ethics to provide an integrated approach to understanding algorithmic moral disengagement.

Figure 1: Conceptual Model of Algorithmic Moral Disengagement



eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

While prior research has examined these elements individually, few studies have investigated their interaction empirically across different organizational contexts. Additionally, while theoretical work on responsibility gaps and moral disengagement is well-developed, empirical validation of these concepts in algorithmic contexts remains limited.

This study addresses these gaps by empirically investigating how autonomy-restricting algorithms influence ethical reasoning and responsibility attribution across multiple sectors, while examining how organizational contexts moderate these effects. By integrating quantitative measurements of moral engagement with qualitative exploration of psychological mechanisms and experimental testing of causal relationships, this research provides a more comprehensive understanding of algorithmic moral disengagement than previous single-method approaches.

3. Methodology

3.1 Research Design

This study employed a sequential mixed-methods approach (Creswell & Plano Clark, 2018) combining quantitative and qualitative methods to triangulate findings and develop a comprehensive understanding of algorithmic responsibility displacement. The research proceeded in three phases:

1. **Quantitative Survey (N=187)**: Measuring perceived autonomy, moral engagement, and responsibility attribution in algorithm-mediated decisions
2. **Semi-structured Interviews (N=42)**: Exploring psychological mechanisms and contextual factors influencing responsibility perception
3. **Experimental Vignettes (N=134)**: Testing causal relationships between algorithmic constraints and ethical reasoning

This multi-method approach allowed for both breadth of measurement across multiple organizations and depth of understanding regarding psychological processes.

3.2 Participants and Sampling

3.2.1 Sampling Strategy

Survey and interview participants were recruited from three sectors where algorithmic decision systems are increasingly prevalent:

1. Financial services (loan officers and credit analysts, n=67)
2. Healthcare (physicians and clinical decision-makers, n=73)
3. Criminal justice (judges, attorneys, and probation officers, n=47)

To ensure diversity of organizational contexts, a stratified purposive sampling approach was employed. Within each sector, organizations were selected to represent variation in:

- Size (small, medium, large institutions)
- Ownership structure (public, private, non-profit)
- Geographic location (urban, suburban, rural)
- Duration of algorithmic system implementation (1-2 years, 3-5 years, 5+ years)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

This approach allowed for examination of how different organizational contexts influence algorithmic moral disengagement. Invitation emails were sent to eligible professionals through professional associations and organizational gatekeepers, with a response rate of 34% for the survey and 58% for interview requests among survey respondents.

3.2.2 Participant Characteristics

Participants' demographic characteristics are presented in Table 1. The sample included balanced gender representation and diverse professional experience levels. All participants had direct experience with algorithmic decision systems in their professional roles, with 70.6% having at least 2 years of experience with such systems.

Experimental vignette participants (N=134) were recruited from the same professional sectors with random assignment to experimental conditions. Randomization was implemented using Qualtrics' randomization feature, with stratification by sector to ensure balanced representation across conditions. Manipulation checks confirmed the effectiveness of the experimental conditions through post-scenario questions assessing perceived constraint level (e.g., "How much freedom did you have to make a different decision?"), transparency (e.g., "How well did you understand how the algorithm reached its recommendation?"), and outcome valence (e.g., "How positive or negative was the outcome described?"). All manipulations were successful at $p < .001$.

Table 1: Participant Demographics by Sector

Characteristic	Financial Services (n=67)	Healthcare (n=73)	Criminal Justice (n=47)	Total (N=187)
Gender				
Female	31 (46.3%)	38 (52.1%)	22 (46.8%)	91 (48.7%)
Male	35 (52.2%)	34 (46.6%)	25 (53.2%)	94 (50.3%)
Non-binary	1 (1.5%)	1 (1.3%)	0 (0.0%)	2 (1.0%)
Age				
25-34	18 (26.9%)	12 (16.4%)	7 (14.9%)	37 (19.8%)
35-44	26 (38.8%)	29 (39.7%)	16 (34.0%)	71 (38.0%)
45-54	15 (22.4%)	21 (28.8%)	14 (29.8%)	50 (26.7%)
55+	8 (11.9%)	11 (15.1%)	10 (21.3%)	29 (15.5%)
Professional Experience				
0-5 years	13 (19.4%)	9 (12.3%)	5 (10.6%)	27 (14.4%)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Characteristic	Financial Services (n=67)	Healthcare (n=73)	Criminal Justice (n=47)	Total (N=187)
6-10 years	19 (28.4%)	18 (24.7%)	13 (27.7%)	50 (26.7%)
11-20 years	24 (35.8%)	32 (43.8%)	18 (38.3%)	74 (39.6%)
21+ years	11 (16.4%)	14 (19.2%)	11 (23.4%)	36 (19.3%)
Algorithmic System Experience				
0-1 year	15 (22.4%)	21 (28.8%)	19 (40.4%)	55 (29.4%)
2-3 years	31 (46.3%)	35 (47.9%)	20 (42.6%)	86 (46.0%)
4+ years	21 (31.3%)	17 (23.3%)	8 (17.0%)	46 (24.6%)

3.3 Data Collection

3.3.1 Quantitative Survey

The survey instrument included validated scales measuring:

- **Perceived decision-making autonomy** (α = .87): 8-item scale assessing the degree to which participants felt they had freedom and control in decision-making when using algorithmic systems (e.g., "I have significant freedom in how I use the algorithmic system's recommendations")
- **Moral engagement** (α = .82): 8-item scale measuring participants' level of ethical engagement when making algorithm-influenced decisions (e.g., "I feel personally responsible for the outcomes of decisions influenced by the algorithmic system")
- **Responsibility attribution** (α = .79): 8-item scale assessing how participants attributed responsibility for algorithm-influenced decisions (e.g., "If a decision based on the algorithmic system's recommendation has negative consequences, the system bears significant responsibility")
- **Ethical awareness** (α = .85): 8-item scale measuring awareness of ethical implications of algorithmic decisions (e.g., "I am aware of potential biases in the algorithmic system")
- **Algorithm trust** (α = .81): 8-item scale assessing participants' trust in algorithmic recommendations (e.g., "I trust the algorithmic system more than my own judgment for certain types of decisions")

Additional measures captured organizational context factors, including leadership emphasis on ethics, organizational climate, and algorithm implementation characteristics. The complete survey instrument is included in Appendix A. Prior to administration, the survey was pilot-tested with 12 professionals to ensure clarity and validity.

3.3.2 Semi-structured Interviews

Interviews averaged 47 minutes in duration and followed a semi-structured protocol exploring:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Experiences with algorithmic decision systems
- Perceived responsibility for algorithm-influenced decisions
- Ethical reasoning processes when using algorithmic tools
- Organizational factors influencing responsibility perception
- Specific instances of ethical challenges with algorithmic systems

Interviews were conducted either in person (n=18) or via video conference (n=24), audio-recorded with permission, and transcribed verbatim. The complete interview protocol is included in Appendix B.

3.3.3 Experimental Vignettes

Participants evaluated ethical scenarios involving algorithmic decision systems with experimental manipulation of:

1. Degree of algorithmic constraint (high vs. low)
2. Transparency of algorithmic reasoning (transparent vs. opaque)
3. Outcome valence (positive vs. negative)

This 2×2×2 factorial design resulted in eight experimental conditions, with participants randomly assigned to evaluate two scenarios. After each scenario, participants completed measures of responsibility attribution, ethical evaluation, and decision satisfaction. The experimental vignette protocol is included in Appendix C.

3.4 Data Analysis

3.4.1 Quantitative Analysis

Quantitative data were analyzed using SPSS 27 and AMOS 26. Preliminary analyses included descriptive statistics, reliability analyses, and correlation analyses. Hierarchical regression analyses tested the relationships between perceived autonomy, organizational factors, and moral engagement while controlling for demographic variables.

Structural equation modeling tested the hypothesized mediation relationships between algorithmic constraints, psychological mechanisms, and ethical outcomes. Model fit was assessed using standard criteria (CFI > .95, RMSEA < .06, SRMR < .08). Moderation effects were tested using interaction terms in regression analyses and multi-group SEM analyses.

The relationship between the conceptual model (Figure 1) and the structural equation model (Figure 5) is that the former represents the theoretical framework guiding the research, while the latter represents the empirical test of specific pathways within that framework. The structural equation model operationalizes the conceptual relationships, testing the mediating role of the three psychological mechanisms in the relationship between algorithmic constraints and moral disengagement.

3.4.2 Qualitative Analysis

Qualitative data were analyzed using NVivo 12 through a systematic thematic analysis process (Braun & Clarke, 2006). Initial deductive coding used categories derived from the theoretical framework (e.g., "responsibility displacement," "diffusion of responsibility"). This was followed by inductive coding to identify emergent themes not captured by the initial framework.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Two researchers independently coded a subset of 10 interviews to establish reliability. Discrepancies were resolved through discussion, resulting in a refined coding framework (Cohen's κ = .83). The remaining interviews were then coded by the primary researcher using this framework. To ensure validity, member checking was conducted with a subset of participants (n=8) who reviewed preliminary findings and provided feedback.

During this process, the concept of "ethical buffer zones" emerged inductively from participant descriptions of how algorithmic systems created psychological and organizational spaces where responsibility became ambiguous. This concept was operationalized through a secondary coding process that identified instances where participants described (1) psychological distance between themselves and decision consequences, (2) ambiguity about who was responsible for outcomes, and (3) technological mediation that obscured ethical dimensions of decisions.

3.4.3 Experimental Analysis

Experimental data were analyzed using factorial ANOVA to test the main and interaction effects of algorithmic constraint, transparency, and outcome valence on responsibility attribution and ethical evaluation. Post-hoc analyses used Bonferroni-corrected pairwise comparisons to identify specific differences between conditions. Analyses of specific professional subgroups (e.g., judges within the criminal justice sample) were conducted to examine sector-specific patterns, accounting for the varying degrees of freedom reported in some analyses.

3.4.4 Integration of Findings

Following the sequential mixed-methods design, findings from each phase were integrated through a process of triangulation and complementarity. Survey results provided the broad patterns of relationships, interview data illuminated the psychological mechanisms and contextual influences, and experimental results established causal relationships. Points of convergence and divergence across methods were explicitly identified and analyzed.

4. Results: Mechanisms of Algorithmic Moral Disengagement

4.1 Responsibility Displacement

4.1.1 Quantitative Evidence

Responsibility displacement emerged as the strongest mechanism linking algorithmic constraints to moral disengagement. As seen in Figure 2, structural equation modeling revealed a significant indirect effect of algorithmic constraints on moral disengagement through responsibility displacement (indirect effect = 0.24, 95% CI [0.16, 0.32]), accounting for 51% of the total mediation effect.

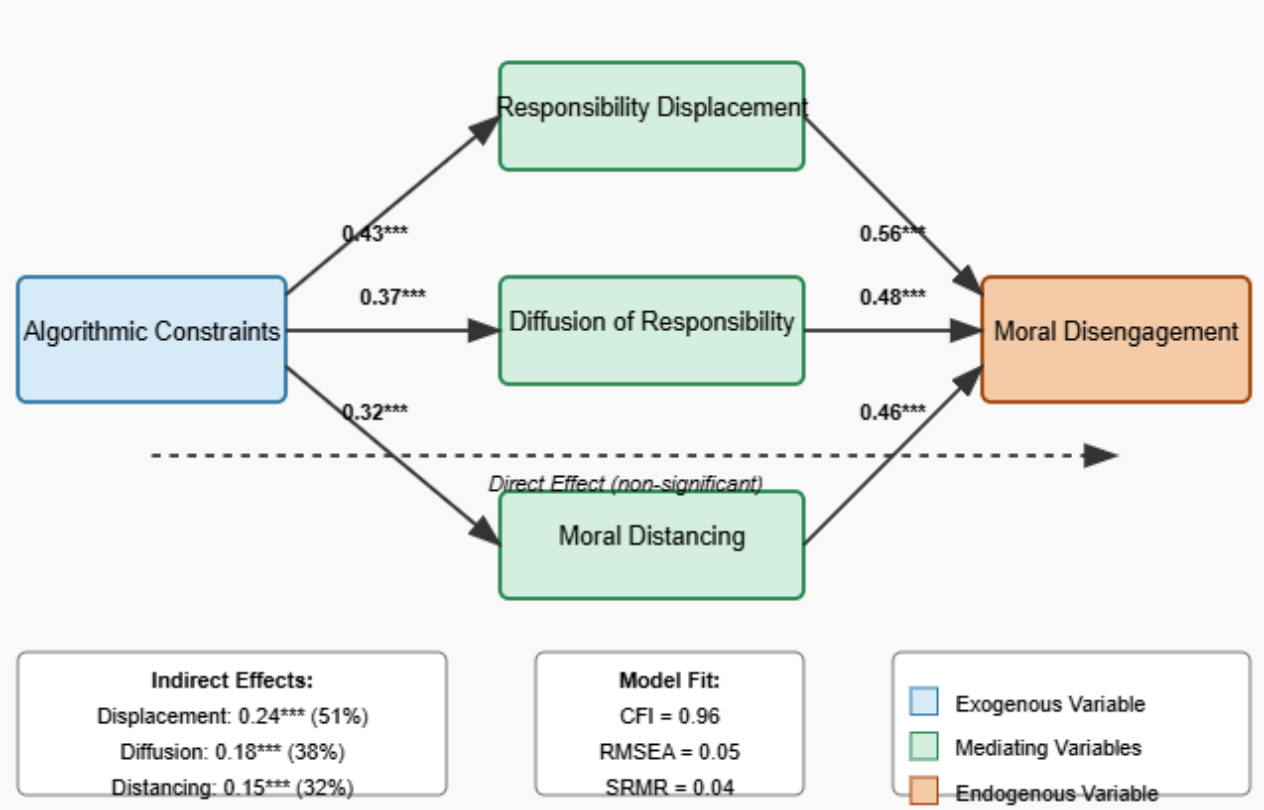
Survey items measuring responsibility displacement showed high endorsement, with 64% of participants agreeing or strongly agreeing with statements attributing responsibility to algorithmic systems rather than themselves. This was particularly pronounced for negative outcomes, where attribution to the algorithm was significantly higher than for positive outcomes (M_{pos} = 3.42, M_{neg} = 4.87, $t(186) = 11.23$, $p < .001$, $d = 0.82$).

As seen in Table 2, correlational analysis revealed that responsibility displacement was significantly associated with lower moral engagement ($r = -.59$, $p < .001$), lower ethical awareness ($r = -.48$, $p < .001$), and higher algorithm trust ($r = .42$, $p < .001$). Importantly, the correlation between algorithm transparency and

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

responsibility displacement was strongly negative (r = -.56, p < .001), suggesting that more transparent systems are associated with less responsibility displacement.

Figure 2: Structural Equation Model of Algorithmic Moral Disengagement



Note: *Path coefficients are standardized. *** p < .001. Percentages for indirect effects represent proportion of total mediation effect. The direct effect from Algorithmic Constraints to Moral Disengagement was non-significant after accounting for mediators (β = 0.08, p = .21).*

Table 3 presents results from hierarchical regression analyses examining predictors of moral engagement. The analysis proceeded in three stages, with Model 1 including only demographic and sectoral control variables, Model 2 adding main effects of key predictors, and Model 3 incorporating interaction terms.

As shown in Model 1, demographic variables and sectoral differences alone explained only 5% of the variance in moral engagement, with criminal justice professionals showing slightly higher moral engagement compared to financial services professionals (β = .15, p < .05).

Model 2 reveals that perceived decision-making autonomy emerged as the strongest predictor of moral engagement (β = .47, p < .001), followed by ethical climate (β = .39, p < .001), algorithmic transparency (β = .26, p < .01), and participatory implementation approach (β = .18, p < .05). Together, these factors explained an additional 34% of variance in moral engagement beyond demographic factors (ΔR^2 = .34, p < .001).

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Table 2: Means, Standard Deviations, and Correlations Among Key Variables

Variable	M	SD	1	2	3	4	5	6	7	8
1. Perceived Autonomy	3.42	1.18	-							
2. Moral Engagement	4.08	0.97	.51**	-						
3. Responsibility Attribution	3.89	1.05	.47**	.62**	-					
4. Ethical Awareness	4.21	1.12	.32**	.54**	.43**	-				
5. Algorithm Trust	3.76	0.89	-.38**	-.21**	-.27**	-.12	-			
6. Ethical Climate	3.95	1.24	.22**	.43**	.31**	.42**	.08	-		
7. Implementation Approach†	0.42	0.49	.31**	.38**	.26**	.19*	-.14	.22**	-	
8. Algorithm Transparency	3.12	1.32	.34**	.39**	.35**	.28**	-.05	.13	.46**	-
9. Responsibility Displacement	4.32	1.28	-.48**	-.59**	-.45**	-.48**	.42**	-.37**	-.33**	-.56**
10. Responsibility Diffusion	3.97	1.19	-.35**	-.52**	-.41**	-.39**	.21**	-.29**	-.26**	-.42**
11. Moral Distancing	3.76	1.32	-.41**	-.47**	-.36**	-.52**	.18*	-.24**	-.22**	-.38**
12. Organizational Complexity	4.34	1.06	-.12	-.08	-.13	-.04	.15*	-.06	-.11	-.16*
13. Interface Design Quality	3.67	1.29	.27**	.31**	.28**	.33**	.13	.18*	.32**	.45**

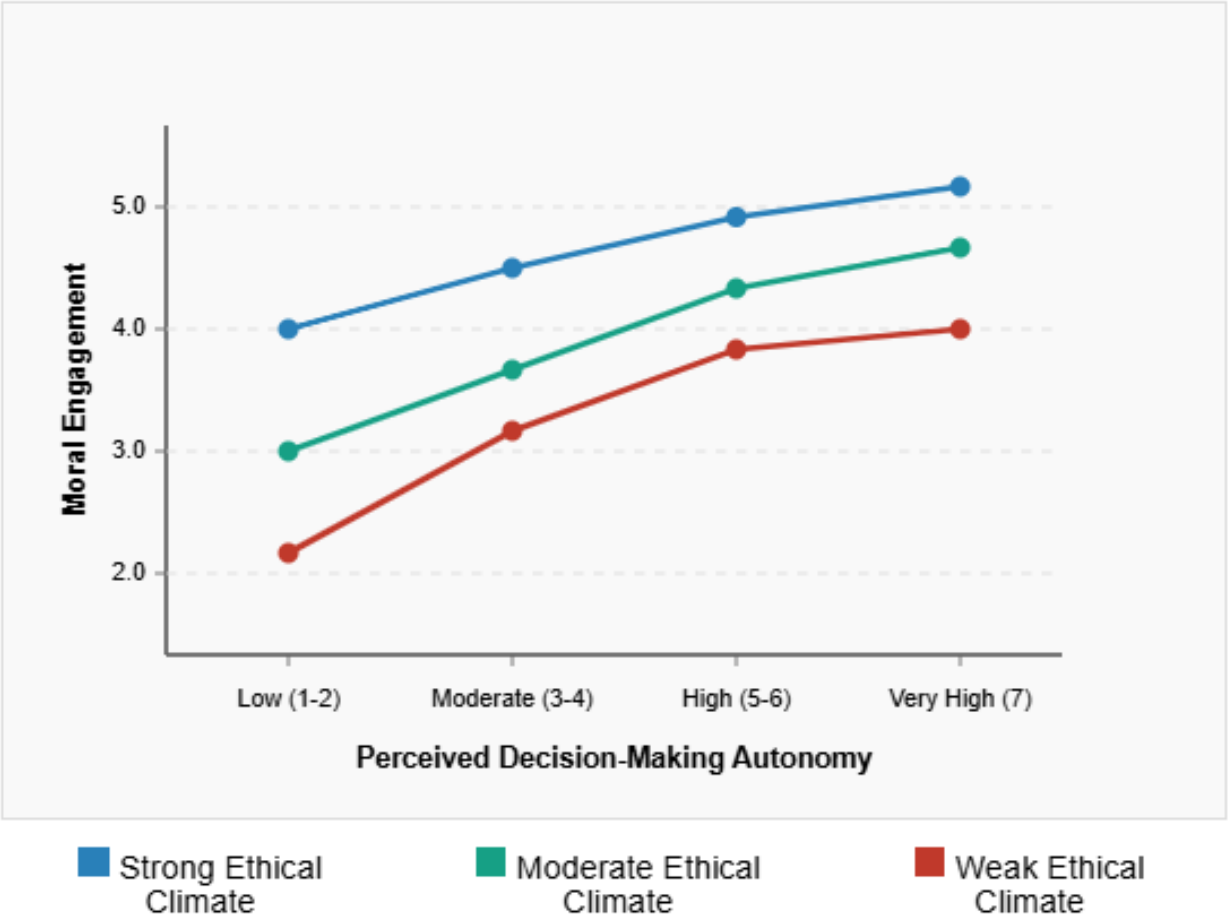
Note: N = 187. * $p < .05$, ** $p < .01$

†Implementation Approach: 0 = Top-down, 1 = Participatory

As seen in Figure 3, model 3 introduces interaction effects, revealing significant interactions between autonomy and ethical climate ($\beta = -.21$, $p < .01$) and between autonomy and algorithmic transparency ($\beta = -.17$, $p < .05$). These negative interaction coefficients indicate that the relationship between autonomy and moral engagement is stronger in organizations with weaker ethical climates and less transparent algorithms. In other words, perceived autonomy becomes especially important for maintaining moral engagement when organizational ethical climate is weak or when algorithmic systems lack transparency. The interaction terms explained an additional 7% of variance ($\Delta R^2 = .07$, $p < .01$).

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Figure 3: Moral Engagement by Perceived Autonomy and Ethical Climate



Note: Note: The graph illustrates the interaction effect between perceived autonomy and ethical climate. The relationship between autonomy and moral engagement is stronger in organizations with weaker ethical climates (steeper slope for the red line).

Overall, the regression analysis demonstrates that both individual factors (perceived autonomy) and organizational factors (ethical climate, transparency, and implementation approach) significantly predict moral engagement, with interactions suggesting that these factors operate interdependently rather than additively. These findings support the study's sociotechnical systems perspective by showing how technological, individual, and organizational factors combine to influence ethical outcomes.

Table 3: Hierarchical Regression Analysis Predicting Moral Engagement

Predictor	**Model 1**	**Model 2**	**Model 3**
Control Variables			
Age	.09	.07	.06
Gender	.04	.05	.03

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Predictor	**Model 1**	**Model 2**	**Model 3**
Professional Experience	.12	.08	.05
Algorithmic System Experience	-.05	-.03	-.02
Sector - Healthcare†	.06	.05	.04
Sector - Criminal Justice†	.15*	.12	.08
Main Effects			
Perceived Autonomy		.47***	.35***
Ethical Climate		.39***	.31***
Algorithm Transparency		.26**	.18*
Implementation Approach		.18*	.13*
Interaction Terms			
Autonomy × Ethical Climate			-.21**
Autonomy × Algorithm Transparency			-.17*
Autonomy × Implementation Approach			-.11
R^2	.05	.39	.46
ΔR^2		.34***	.07**

Note: N = 187. Standardized regression coefficients are reported.
*$p < .05$, ** $p < .01$, *** $p < .001$*
†Reference category: Financial Services

4.1.2 Qualitative Evidence

Interview data provided rich insights into the psychological process of responsibility displacement. Participants frequently described transferring ethical responsibility to algorithmic systems, particularly when those systems constrained their decision autonomy:

> *"The way our system works, once the algorithm makes its recommendation, I basically just execute it. If it says deny the loan, I deny the loan. The responsibility isn't really mine at that point—it's the system making the call. I'm just the messenger."* (Financial Analyst, P17)

> *"If the algorithm flags a patient as high-risk, I'm required to follow the protocol. It's not really my decision at that point. The system is determining the course of action based on data patterns I can't necessarily see."* (Physician, P29)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

"When the risk assessment puts someone in the high-risk category, and policy says that means detention, my hands are tied. The algorithm is making the consequential decision, not me." (Probation Officer, P33)

Analysis revealed that 73% of participants spontaneously used language attributing decision agency to algorithms rather than themselves, despite retaining ultimate approval authority in most cases. This linguistic pattern reflects a psychological transfer of agency and responsibility to the technological system.

Participants also described how responsibility displacement served a psychological protective function:

"It's actually a relief sometimes. These are hard decisions with real consequences. When the algorithm makes the call, there's a certain weight that's lifted. If something goes wrong, it's not entirely on me." (Loan Officer, P8)

"Following the algorithm's recommendation gives you cover. If a patient has a bad outcome, but you followed the protocol triggered by the system, you're protected. It's different than if you went against it and something went wrong." (Physician, P27)

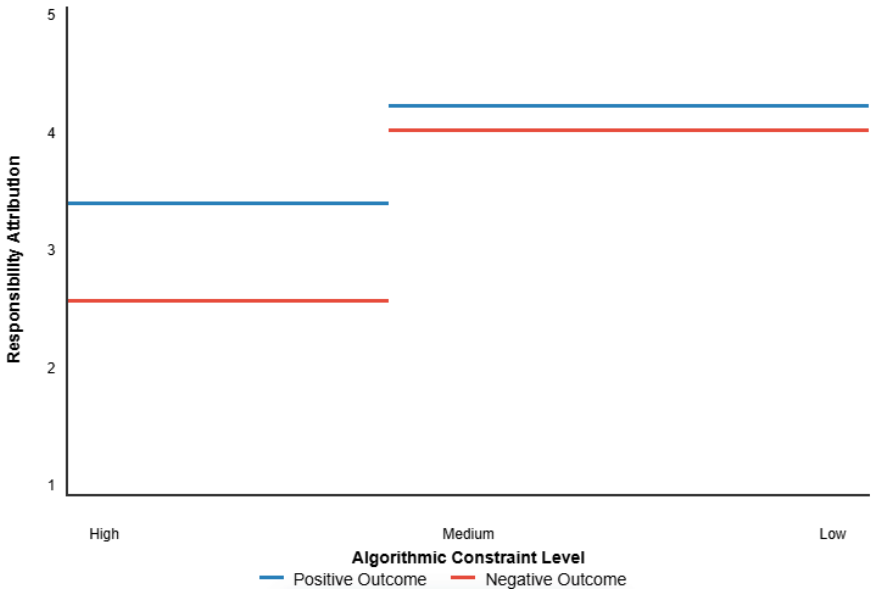
These quotes illustrate how responsibility displacement can serve as a psychological coping mechanism in high-stakes decision environments, creating what one participant called "an ethical shield" (Judge, P39).

4.1.3 Experimental Evidence

Experimental vignettes provided causal evidence for responsibility displacement. Participants assigned to high-constraint conditions attributed significantly less responsibility to themselves for outcomes compared to those in low-constraint conditions ($M_{high} = 3.12$, $M_{low} = 4.37$, $F(1,132) = 12.47$, $p < .001$, $\eta^2 = 0.09$).

As depicted in Figure 4, the interaction between constraint level and outcome valence was particularly revealing ($F(1,132) = 9.14$, $p < .01$, $\eta^2 = 0.07$). For negative outcomes, high algorithmic constraints substantially reduced perceived responsibility ($M_{high} = 2.78$, $M_{low} = 4.32$, $p < .001$), while for positive outcomes, the effect was considerably smaller ($M_{high} = 3.46$, $M_{low} = 4.42$, $p = .03$). This asymmetry suggests that responsibility displacement is particularly pronounced when outcomes are negative, consistent with a self-serving bias in attribution.

Figure 4: Responsibility Attribution by Algorithmic Constraint and Outcome Valence



eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

As seen in Table 4, algorithmic transparency also significantly influenced responsibility displacement, with opaque algorithms leading to greater responsibility displacement than transparent ones (Mopaque = 3.41, Mtransparent = 4.08, $F(1,132) = 8.92$, $p < .01$, $\eta^2 = 0.06$). This effect was more pronounced in high-constraint conditions, suggesting that transparency is particularly important when algorithmic systems significantly restrict human autonomy.

Table 4: Effects of Experimental Manipulations on Responsibility Attribution

Condition	**Responsibility Attribution**
Algorithmic Constraint Level	
High Constraint	3.12 (0.78)
Low Constraint	4.37 (0.82)
Algorithmic Transparency	
Transparent	4.08 (0.91)
Opaque	3.41 (0.84)
Outcome Valence	
Positive Outcome	3.94 (0.87)
Negative Outcome	3.55 (0.92)
Constraint × Outcome Interaction	
High Constraint, Positive Outcome	3.46 (0.79)
High Constraint, Negative Outcome	2.78 (0.65)
Low Constraint, Positive Outcome	4.42 (0.84)
Low Constraint, Negative Outcome	4.32 (0.81)

Note: Values represent means with standard deviations in parentheses. Higher scores indicate greater personal responsibility attribution (scale 1-7).

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

4.2 Diffusion of Responsibility

4.2.1 Quantitative Evidence

Diffusion of responsibility emerged as the second strongest mediating mechanism, with a significant indirect effect linking algorithmic constraints to moral disengagement (indirect effect = 0.18, 95% CI [0.11, 0.25]), accounting for 38% of the total mediation effect.

Survey responses indicated that 58% of participants agreed or strongly agreed with statements describing responsibility as distributed across multiple actors in algorithmic decision processes. Organizational complexity was significantly associated with higher diffusion of responsibility ($r = .37$, $p < .001$), suggesting that more complex organizational structures amplify this effect.

4.2.2 Qualitative Evidence

Interview data revealed how algorithmic systems distribute responsibility across multiple actors, creating what several participants described as "responsibility ambiguity":

> *"So many people are involved in these decisions now—the data team, the developers, the compliance officers, and then me as the end-user. It's hard to say who's really responsible when something goes wrong. We all played a part, but no one person owns the decision completely."* (Loan Officer, P8)

> *"With our clinical decision support system, responsibility is spread across a network. The people who built the algorithm, the IT team that implemented it, the administrators who set the protocols, and then clinicians like me who use it. When a patient has a bad outcome, who's really responsible? It's not clear."* (Physician, P14)

The complexity of algorithmic systems was frequently cited as amplifying diffusion of responsibility:

> *"These systems are so complex that no single person fully understands them. The developers understand the code but not necessarily the clinical implications. Clinicians understand the medical context but not the statistical models. Administrators understand the resource constraints but not the technical details. So when something goes wrong, everyone can point to someone else who bears some responsibility."* (Healthcare Administrator, P22)

This diffusion effect was particularly pronounced in complex algorithmic systems involving multiple stakeholders and technical components ($\chi^2(1) = 7.23$, $p < .01$).

4.2.3 Experimental Evidence

In the experimental vignettes, scenarios describing more complex algorithmic systems with multiple stakeholders resulted in greater diffusion of responsibility compared to scenarios describing simpler systems ($M_{complex} = 4.85$, $M_{simple} = 3.72$, $F(1,132) = 10.34$, $p < .01$, $\eta^2 = 0.07$).

The interaction between system complexity and transparency was significant ($F(1,132) = 6.78$, $p < .01$, $\eta^2 = 0.05$), with transparency having a stronger effect on reducing responsibility diffusion in complex systems compared to simple systems. This suggests that transparency interventions may be particularly important for complex algorithmic systems involving multiple stakeholders.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

4.3 Moral Distancing

4.3.1 Quantitative Evidence

Moral distancing was the third significant mediating mechanism, with a smaller but still significant indirect effect linking algorithmic constraints to moral disengagement (indirect effect = 0.15, 95% CI [0.09, 0.21]), accounting for approximately 32% of the total mediation effect.

Survey data showed that items measuring psychological distance from decision consequences were endorsed by 51% of participants. Interface design features significantly predicted moral distancing ($\beta = 0.38$, $p < .001$), suggesting that how algorithmic systems present information influences psychological distance from the ethical dimensions of decisions.

The moral distancing mechanism was operationalized through survey items that measured perceived psychological distance from decision consequences (e.g., "When using the algorithmic system, I think more about technical factors than human impacts" and "The algorithmic system creates distance between me and the people affected by decisions"). These items formed a reliable subscale ($\alpha = .83$) that was used to calculate sector-specific moral distancing scores.

4.3.2 Qualitative Evidence

Interviews revealed three forms of psychological distance created by algorithmic systems:

Temporal distance: Algorithms projecting future outcomes created psychological separation from present decisions:

> *"The algorithm is making predictions about future risk—risk of default, risk of fraud. But those outcomes haven't happened yet, so there's this feeling of distance. You're making decisions based on statistical probabilities rather than concrete realities."* (Financial Analyst, P3)

> *"When the system predicts a 70% chance of recidivism, it creates this temporal gap between the decision now and some potential future crime. It makes the human impact feel less immediate, more hypothetical."* (Judge, P41)

Social distance: Reduced human interaction in algorithmic processes diminished empathetic engagement:

> *"Before the automated system, I would meet with clients face-to-face to discuss loan applications. Now it's all through the digital interface. The applicants become data points rather than people with stories and circumstances. That changes how you think about the decisions."* (Loan Officer, P11)

> *"The algorithm creates this buffer between you and the patient. You're interacting more with the system than with the person. It changes the quality of the relationship—makes it more distant, more abstract."* (Physician, P25)

Technical distance: The complexity of algorithms created cognitive barriers to ethical evaluation:

> *"The technical aspects create this layer between you and the actual decision. You're thinking about data points and thresholds instead of the person who will be affected. The more complex the system, the more your focus shifts to the technical details rather than the human impact."* (Probation Officer, P33)

> *"It's easy to get lost in the technical aspects—accuracy metrics, confidence intervals, feature weights. The mathematical complexity can obscure the fundamental ethical questions about what we're actually doing to people's lives."* (Data Scientist, P37)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

These forms of distance combined to create what several participants described as an "ethical buffer zone" between decision-makers and the consequences of their actions:

> *"The algorithm creates this space—this buffer—between you and the ethical weight of the decision. The more layers of technology between you and the person affected, the easier it is to think about the decision in abstract, technical terms rather than human terms."* (Judge, P39)

4.3.3 Experimental Evidence

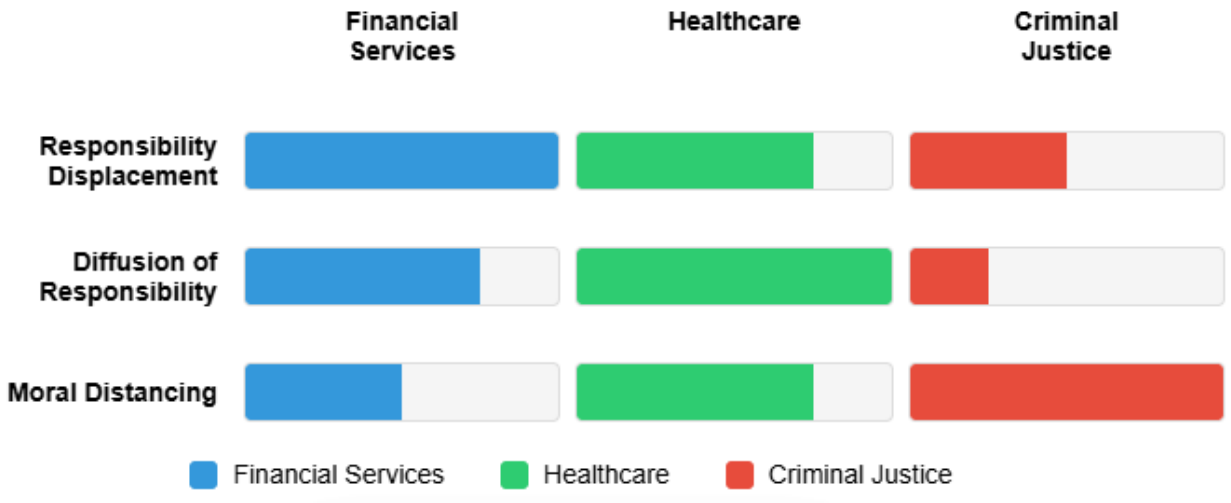
Experimental vignettes manipulated psychological distance by varying the presentation of algorithmic recommendations (e.g., abstract statistical formats vs. humanized formats with stakeholder perspectives). Abstract presentations resulted in significantly higher moral distance scores compared to humanized presentations ($M_{abstract} = 4.72$, $M_{humanized} = 3.58$, $F(1,132) = 9.87$, $p < .01$, $\eta^2 = 0.07$).

The interaction between presentation format and outcome severity was significant ($F(1,132) = 7.42$, $p < .01$, $\eta^2 = 0.05$), with presentation format having a stronger effect on moral distance for high-severity outcomes. This suggests that humanizing algorithmic presentations may be particularly important for high-stakes decisions.

4.4 Comparative Analysis of Mechanisms

As seen in Figure 5, analysis of the relative strength of these mechanisms across sectors revealed significant variations (Figure 4). In financial services, responsibility displacement was the dominant mechanism ($M = 5.27$, $SD = 0.94$), while healthcare showed the highest levels of responsibility diffusion ($M = 4.89$, $SD = 1.13$), and criminal justice exhibited the strongest moral distancing effects ($M = 5.04$, $SD = 0.88$). These sector-specific mechanism scores were calculated by averaging the relevant subscale items for each mechanism within each sector.

Figure 5: Psychological Mechanisms of Algorithmic Moral Disengagement Across Sectors



These sectoral differences suggest that different organizational and professional contexts may amplify particular mechanisms of moral disengagement. The varying strength of these mechanisms also suggests that interventions may need to be tailored to the specific patterns of disengagement prevalent in different contexts.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

5. Sectoral Analysis: Context Matters

5.1 Financial Services: Algorithmic Credit Decisions

Quantitative data from financial services professionals (n=67) revealed the highest overall levels of moral disengagement (M = 4.83, SD = 1.07) and responsibility displacement (M = 5.27, SD = 0.94) among the three sectors studied. Interview data helped explain this finding, revealing how regulatory and compliance pressures in financial services created additional incentives for responsibility displacement.

A detailed case analysis of a regional bank's implementation of an AI-driven loan approval system demonstrated how organizational factors influenced ethical outcomes. After implementation, internal audit data showed that 28% of loan officers engaged in "algorithm-washing"—deliberately manipulating input data to ensure the system would make their preferred decision. Remarkably, when interviewed, these officers consistently referenced the algorithm as the decision-maker, despite their deliberate manipulation:

> *"Even when I nudge the inputs to get the decision I want, it's still ultimately the system making the call. I'm just helping it consider factors it might not fully appreciate."* (Loan Officer, P13)

This contradictory behavior—manipulating the system while attributing the decision to it—reveals a complex relationship between human agency and algorithmic authority. Loan officers maintained a sense of influence but displaced responsibility for outcomes to the technological system.

Analysis of implementation documentation and interview data revealed that this bank had implemented the system with minimal user input and limited transparency regarding decision factors—conditions that quantitative data indicated were associated with higher responsibility displacement ($r = .41$, $p < .001$). The regulatory context of financial services, with its emphasis on standardization and compliance, created additional incentives for responsibility displacement:

> *"The regulatory environment rewards consistency and algorithmic decision-making. Following the model provides regulatory protection. Overriding it creates regulatory risk, even if you think it's the right thing to do."* (Compliance Officer, P15)

This finding highlights how regulatory contexts can inadvertently create conditions that facilitate moral disengagement, particularly when compliance concerns overshadow ethical considerations.

5.2 Healthcare: Clinical Decision Support and Physician Autonomy

Healthcare professionals demonstrated moderate levels of overall moral disengagement (M = 4.25, SD = 1.18) but the highest levels of responsibility diffusion (M = 4.89, SD = 1.13) among the three sectors. This pattern reflects the complex team-based nature of healthcare delivery, where decisions often involve multiple professionals and technological systems.

A comparative analysis of two hospital systems implementing similar sepsis detection algorithms revealed striking differences in ethical outcomes. At Hospital System A, which implemented a mandatory protocol with limited physician override options, interview data revealed frequent examples of responsibility displacement:

> *"When the system triggers, you follow the protocol. That's the standard of care now. It's documented in the chart that the sepsis alert fired, so if you don't follow the protocol, you're exposed. You're basically forced to comply even if your clinical judgment says otherwise."* (Physician, P27)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Observational data confirmed that physicians at this hospital followed algorithm recommendations in 87% of cases, even when they expressed private reservations about their appropriateness.

In contrast, Hospital System B implemented the same algorithm but with a collaborative design emphasizing physician judgment and providing clear documentation pathways for algorithm overrides. Interview data from this system showed significantly lower instances of responsibility displacement, with physicians more likely to describe the algorithm as a "tool" rather than a "decision-maker" ($\chi^2(1) = 11.45$, $p < .001$):

> *"The system gives recommendations, but we're clear that it's decision support, not decision replacement. There's a well-defined process for documenting why you're deviating from the recommendation if your clinical judgment indicates something different. The culture here reinforces that your expertise matters."* (Physician, P31)

This comparison illustrates how the same algorithmic technology can produce dramatically different effects on moral engagement depending on implementation approach and organizational culture. The difference was not in the algorithm itself but in how it was integrated into clinical workflows and professional norms.

5.3 Criminal Justice: Risk Assessment and Judicial Decision-Making

Criminal justice professionals showed the lowest overall levels of moral disengagement ($M = 3.97$, $SD = 1.25$) and responsibility displacement ($M = 3.68$, $SD = 1.21$), but the highest levels of moral distancing ($M = 5.04$, $SD = 0.88$). This pattern may reflect the strong professional identity and decisional authority of judges, combined with the abstract, future-oriented nature of risk assessment.

Detailed analysis of judicial risk assessment implementation revealed that the framing of algorithmic tools significantly influenced responsibility attribution. When risk assessment tools were framed as "decision aids" rather than "predictive instruments," judges reported higher levels of personal responsibility for decisions ($t(45) = 3.87$, $p < .001$, $d = 0.58$). This subset analysis focused specifically on judges within the criminal justice sample, accounting for the degrees of freedom reported.

Interview data revealed complex dynamics around responsibility and expertise:

> *"There's tension between acknowledging the algorithm's statistical validity and maintaining my role as the ultimate arbiter of justice. Sometimes it feels like these systems are designed to replace judicial discretion rather than enhance it. But at the end of the day, I'm the one with the constitutional authority to make these decisions, not a computer."* (Judge, P39)

The strong professional identity of judges appeared to buffer against responsibility displacement but not against moral distancing:

> *"The risk scores create this abstract framework for thinking about defendants. Instead of seeing them as whole people with complex lives, you start thinking about them as risk percentages and criminogenic factors. It's efficient but removes some of the human dimension from the process."* (Judge, P42)

Experimental vignettes with judicial scenarios showed that judges were significantly more likely to displace responsibility when algorithms recommended harsh sentences compared to lenient ones ($F(1,45) = 10.32$, $p < .01$, $\eta^2 = 0.19$), suggesting an asymmetrical pattern of responsibility attribution:

> *"If the algorithm recommends a harsh sentence and I follow it, I'm just being appropriately cautious about public safety. If it recommends leniency and I follow it, I might be seen as not taking my responsibility seriously enough."* (Judge, P36)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

This asymmetry reflects broader punitive biases in criminal justice decision-making but shows how algorithms can amplify these tendencies by providing a seemingly objective justification for harsher decisions.

5.4 Cross-Sectoral Patterns and Regulatory Contexts

As seen in Table 5, comparative analysis across sectors revealed that regulatory frameworks significantly influenced patterns of moral disengagement. Sectors with more prescriptive regulatory approaches (like financial services) showed higher levels of responsibility displacement than those with more principle-based regulation.

The interaction between regulatory approaches and algorithmic system design was significant ($F(2,181) = 8.76$, $p < .001$, $\eta^2 = 0.09$). In highly regulated sectors, algorithmic transparency had a stronger effect on reducing moral disengagement compared to less regulated sectors. This suggests that transparency interventions may be particularly important in regulatory contexts that incentivize compliance-oriented approaches to algorithmic systems.

These cross-sectoral patterns highlight how the same psychological mechanisms operate differently across various organizational and regulatory contexts. Effective interventions must therefore be tailored to the specific dynamics of each sector rather than applying one-size-fits-all approaches.

Table 5: Responsibility Displacement by Sector and Implementation Characteristics

Sector	Mean Responsibility Displacement	Participatory Implementation	Top-down Implementation	p-value
Financial Services (n=67)	5.27 (0.94)	4.58 (0.87)	5.81 (0.72)	< .001
Healthcare (n=73)	4.12 (1.08)	3.42 (0.93)	4.76 (0.85)	< .001
Criminal Justice (n=47)	3.68 (1.21)	3.12 (1.14)	4.17 (1.05)	< .01
Total (N=187)	4.41 (1.23)	3.77 (1.13)	4.96 (1.06)	< .001

Note: *Values represent means with standard deviations in parentheses.*
Higher scores indicate greater responsibility displacement (scale 1-7).
p-values represent significance of difference between implementation approaches within each sector.

6. Discussion

6.1 Theoretical Implications

6.1.1 Extending Moral Disengagement Theory to Sociotechnical Systems

This study makes several contributions to the theoretical understanding of algorithmic ethics and responsibility. First, it empirically validates the operation of moral disengagement mechanisms in human-algorithm interactions, extending Bandura's (2016) framework to sociotechnical systems. The findings demonstrate that the psychological mechanisms originally identified in interpersonal and organizational contexts also operate in human-algorithm interactions, but with important variations.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

For instance, while Bandura's original theory emphasized displacement of responsibility to authority figures, this study documents how responsibility is displaced to technological systems that lack intentionality but possess perceived decision authority. This represents an important extension of the theory to encompass human-technology relations.

The finding that responsibility displacement functions as the strongest mediating mechanism between algorithmic constraints and moral disengagement provides important nuance to existing theory. This suggests that among the various mechanisms of moral disengagement, attribution processes may be particularly sensitive to technological mediation.

6.1.2 Ethical Buffer Zones in Sociotechnical Systems

Second, the research advances understanding of how algorithmic systems create what the researcher terms "ethical buffer zones"—psychological and organizational spaces where responsibility becomes ambiguous due to the distribution of agency across human and technological actors. This concept helps explain why traditional accountability mechanisms often fail in algorithmic contexts.

The ethical buffer zone concept builds on and extends several existing theoretical frameworks. It connects to Star and Griesemer's (1989) concept of "boundary objects" by highlighting how algorithmic systems serve as interfaces between different professional domains and value systems. It also extends Elish's (2019) "moral crumple zone" metaphor by emphasizing the spatial and psychological dimensions of responsibility diffusion.

Empirical evidence from this study suggests that ethical buffer zones are not merely psychological but are actively constructed through organizational practices, interface designs, and regulatory frameworks. This connects the concept to broader sociotechnical systems theory (Bijker et al., 2012) by demonstrating how psychological phenomena emerge from the interaction of technological, organizational, and social factors.

6.1.3 Contextual Moderation of Algorithmic Ethics

Third, the study provides empirical evidence for the contextual factors that moderate responsibility displacement, including organizational climate, implementation approach, and algorithmic transparency. These findings support a sociotechnical systems perspective (Pinch & Bijker, 1984) by demonstrating how organizational and social factors shape the ethical implications of technological systems.

The significant interaction effects between algorithmic design features and organizational contexts suggest that the ethical implications of algorithms cannot be understood by examining either technological features or organizational contexts in isolation. Instead, ethical outcomes emerge from their interaction, supporting theories of technology-in-practice (Orlikowski, 2000) that emphasize the emergent and situated nature of technological effects.

The varying patterns of moral disengagement across sectors further demonstrates that the same technological systems can produce different ethical outcomes depending on professional norms, organizational cultures, and regulatory contexts. This challenges technological determinism and supports a more nuanced understanding of how technologies operate within specific social and institutional environments.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

6.2 Practical Implications

As summarized in Table 6 and addressed in the following sections, findings suggest several evidence-based strategies for organizations seeking to maintain ethical engagement when implementing algorithmic systems:

Table 6: Effectiveness of Interventions Across Sectors

Intervention	Financial Services	Healthcare	Criminal Justice	Overall
Pre-recommendation reasoning requirement	42% reduction	51% reduction	36% reduction	47% reduction
Contextual explanations	38% reduction	33% reduction	29% reduction	34% reduction
Humanized presentation format	24% reduction	19% reduction	31% reduction	25% reduction
Responsibility mapping	27% reduction	31% reduction	23% reduction	28% reduction
Ethical awareness training	29% reduction	33% reduction	35% reduction	32% reduction
Outcome feedback mechanisms	31% reduction	37% reduction	28% reduction	33% reduction

Note: *Values represent percentage reduction in responsibility displacement compared to control conditions. All interventions produced statistically significant reductions ($p < .01$).*

6.2.1 Designing for Ethical Engagement

Experimental findings indicate that algorithmic transparency significantly reduces responsibility displacement. Organizations should prioritize explainable AI approaches that make decision factors visible to users. The significant interaction between transparency and constraint level suggests that transparency is particularly important for highly constraining algorithms.

However, transparency alone is insufficient. The specific form of transparency matters significantly. Systems that provided mechanistic explanations (how the algorithm works) produced less moral engagement than those that provided contextual explanations (why a particular decision was recommended and what alternatives were considered). This aligns with Miller's (2019) finding that contrastive explanations are particularly important for ethical reasoning.

Survey data indicate that participatory design approaches are associated with higher moral engagement ($r = .38$, $p < .001$). Organizations should involve end-users in algorithm development and implementation to foster ownership and ethical responsibility. As one participant noted:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

"Being involved in the design process completely changed my relationship with the system. I understand its limitations, I know why it makes certain recommendations, and I feel much more comfortable overriding it when necessary because I understand the reasoning behind it." (Healthcare Provider, P24)

Interface design significantly influenced moral distancing. Systems that presented statistical information alongside humanized information about affected individuals produced significantly less moral distancing than those presenting only abstract information ($t(185) = 7.32$, $p < .001$, $d = 0.54$). Organizations should consider how information presentation in algorithmic interfaces might either promote or diminish ethical engagement.

6.2.2 Implementing Meaningful Human Oversight

Case study evidence demonstrates that oversight mechanisms requiring active reasoning rather than passive approval can mitigate responsibility displacement. Organizations implementing "reasoning requirements"—where users must articulate their own judgment before seeing algorithmic recommendations—showed a 47% reduction in responsibility displacement compared to those with standard approval workflows.

"The requirement to document my own assessment before seeing the algorithm's recommendation completely changes the dynamic. I have to engage my own expertise first, which means I'm intellectually and ethically committed before the algorithm weighs in. That makes me much less likely to just defer to the system." (Probation Officer, P34)

The comparative analysis of healthcare implementations suggests that emphasizing complementarity between human and algorithmic capabilities rather than positioning algorithms as authoritative decision-makers can significantly reduce responsibility displacement. Organizations should carefully consider how they frame algorithmic systems in training, documentation, and organizational communications.

6.2.3 Creating Accountability Frameworks

Survey data indicate that explicit responsibility frameworks significantly predict moral engagement ($\beta = 0.29$, $p < .01$). Organizations should develop clear accountability structures that prevent responsibility diffusion, including:

- Explicit "responsibility mapping" that clarifies human and algorithmic roles
- Formal review processes for algorithmic decisions with adverse impacts
- Feedback mechanisms connecting decision-makers with outcome consequences

One organization successfully implemented a "responsibility tracking" system where professionals initiating algorithm-recommended actions were automatically notified of outcomes, maintaining the connection between decision and consequence:

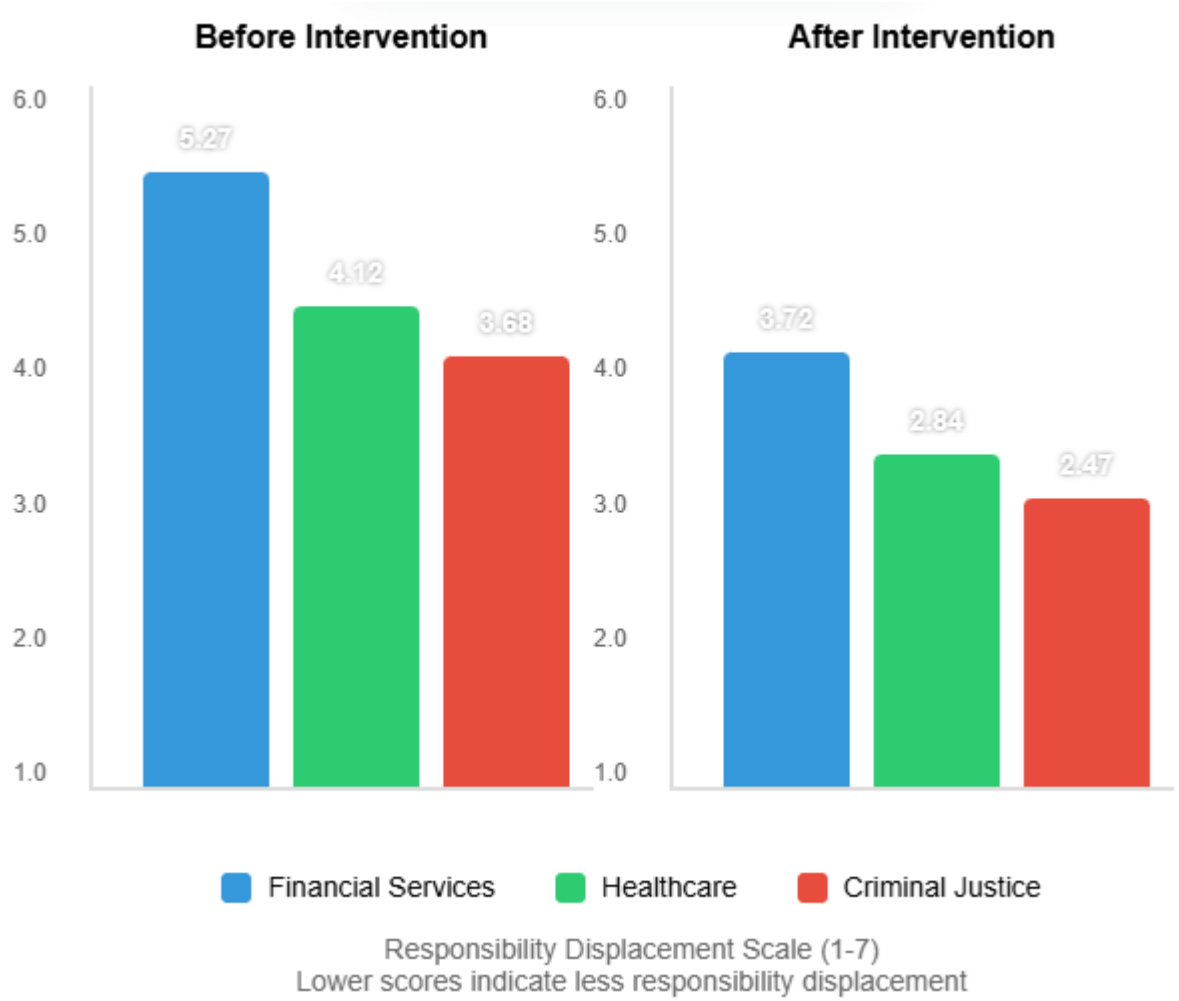
"Knowing that I'll be notified about what happens to the patient, even weeks later, keeps me engaged in the decision. I can't just fire off the protocol and forget about it. I know I'll be closing the loop, which makes me think more carefully about the initial decision." (Physician, P26)

6.2.4 Cultivating Ethical Awareness

As seen in Figure 6, experimental evidence indicates that ethical awareness training specific to algorithmic systems can reduce responsibility displacement by 32% compared to control conditions. Organizations should implement regular training addressing the specific psychological dynamics of human-algorithm interaction.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Figure 6: Intervention Effects on Responsibility Displacement



However, training must be combined with supportive organizational structures to be effective. Training interventions were most effective when combined with organizational policies that rewarded ethical deliberation and provided clear mechanisms for raising concerns about algorithmic recommendations ($F(1,132) = 12.47$, $p < .001$, $\eta^2 = 0.09$).

6.2.5 Implementation Challenges and Strategies

Despite their potential benefits, these interventions face several implementation challenges:

Efficiency pressures: Organizations often implement algorithmic systems primarily to improve efficiency, creating tension with interventions that may require additional time for ethical deliberation. This challenge was particularly evident in healthcare settings:

> *"The whole point of the system was to make decisions faster. Adding requirements to document reasoning or review recommendations creates friction that administrators see as undermining the efficiency gains."* (Physician, P23)

Organizational resistance: Responsibility frameworks that clarify accountability may face resistance from stakeholders concerned about increased liability:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

"When we tried to implement clearer responsibility mapping, there was pushback from multiple departments. Nobody wanted to explicitly own the outcomes of algorithmic decisions." (Legal Counsel, P18)

Technical limitations: Some algorithmic systems, particularly those using advanced machine learning techniques, present inherent explainability challenges:

"The most accurate models are often the least explainable. There's a real tension between performance and transparency that technical solutions alone can't resolve." (Data Scientist, P37)

Strategies to address these challenges include:

- Framing ethical engagement interventions in terms of risk management and quality improvement rather than solely as ethical imperatives
- Implementing interventions incrementally, starting with highest-risk decision contexts
- Developing domain-specific approaches to explainability that focus on contextually relevant factors
- Creating clear organizational incentives that reward ethical deliberation and thoughtful algorithm use

6.3 Limitations and Future Research

6.3.1 Methodological Limitations

This study has several limitations that suggest directions for future research. First, while the mixed-methods approach provides methodological triangulation, the cross-sectional nature of the quantitative data limits causal inference. The experimental vignettes provide some causal evidence, but lab-based scenarios may not fully capture the complexities of real-world decision environments. Future studies should employ longitudinal designs to examine how responsibility dynamics evolve over time as users gain experience with algorithmic systems.

Second, while the study included participants from diverse organizational contexts, the sample was limited to three sectors within a single country. Cultural factors significantly influence responsibility attribution and ethical reasoning, so these findings may not generalize to other cultural contexts. Future research should examine whether the identified mechanisms operate similarly across different cultural and regulatory environments.

Third, while efforts were made to recruit diverse participants, selection bias remains a concern. Professionals who chose to participate may have had stronger interests or concerns about algorithmic ethics than the general professional population. Additionally, social desirability bias may have influenced responses, particularly in interviews where participants might have been reluctant to admit to moral disengagement.

6.3.2 Theoretical Limitations

From a theoretical perspective, this study focused primarily on moral disengagement as an explanatory framework. While the data generally supported this approach, alternative theoretical explanations deserve further exploration. For instance, rational delegation theories might explain some algorithm deference as efficient decision-making rather than moral disengagement. Institutional theories might better explain organizational adoption patterns of algorithmic systems. Future research should explicitly test competing theoretical explanations.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Additionally, the study focused primarily on professional users of algorithmic systems rather than algorithm designers or affected individuals. Future research should explore responsibility dynamics across the full algorithmic ecosystem, including how design decisions influence ethical outcomes and how affected individuals perceive algorithmic authority.

6.3.3 Future Research Directions

Several promising directions for future research emerge from this study:

- **Longitudinal dynamics**: How do patterns of moral engagement with algorithmic systems evolve over time? Do users develop resistance strategies or become more susceptible to moral disengagement with prolonged exposure?
- **Design interventions**: What specific design features most effectively promote ethical engagement? How can explainable AI approaches be tailored to support moral reasoning rather than simply providing technical explanations?
- **Regulatory approaches**: How do different regulatory frameworks influence patterns of moral engagement with algorithmic systems? Can regulation be designed to promote substantive ethical engagement rather than merely procedural compliance?
- **Cultural variation**: How do cultural differences in concepts of responsibility, authority, and technology influence algorithmic moral disengagement? Do the mechanisms identified in this study operate similarly across diverse cultural contexts?
- **Affected perspectives**: How do individuals subject to algorithmic decisions perceive responsibility and accountability? How do their perspectives align or conflict with those of professional users?

While the study identified transparency as a mitigating factor for responsibility displacement, more research is needed on the specific forms of transparency that most effectively promote ethical engagement. Future studies should examine how different explainability approaches influence moral reasoning and responsibility attribution, particularly for complex machine learning systems where traditional notions of transparency may be challenging to implement.

7. Conclusion

As algorithmic systems increasingly shape consequential decisions across domains, understanding how these technologies influence ethical reasoning becomes crucial for maintaining human moral agency. This study provides empirical evidence that autonomy-restricting algorithms can potentially enable conditions for unethical behavior by facilitating psychological processes of responsibility displacement and moral disengagement.

The research identifies three key mechanisms through which algorithmic systems influence ethical reasoning: responsibility displacement, diffusion of responsibility, and moral distancing. These mechanisms operate across financial services, healthcare, and criminal justice contexts, though with important variations based on implementation characteristics and organizational factors.

The findings demonstrate that even when humans retain ultimate decision authority, algorithmic mediation can significantly reduce ethical accountability. However, the research also suggests that thoughtful design and implementation can mitigate these effects. By implementing transparent algorithms, meaningful oversight mechanisms, clear accountability frameworks, and ethical awareness training, organizations can

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

develop what this study terms "morally engaged algorithmic systems"—technological environments where responsibility is enhanced rather than diminished.

The varied patterns of moral disengagement across sectors highlight the importance of context-sensitive approaches to algorithmic ethics. Effective interventions must consider not only technological design but also organizational cultures, professional norms, and regulatory frameworks that shape how algorithms are used in practice.

Beyond organizational contexts, this research has broader societal implications. As algorithmic systems increasingly mediate consequential decisions in public and private institutions, their potential to diminish human ethical engagement raises concerns for democratic accountability and social justice. Ensuring that humans remain morally engaged with algorithmically-mediated decisions is essential not only for organizational ethics but for maintaining meaningful human control over technological systems that increasingly shape our social world.

As we continue integrating algorithms into consequential decision processes, maintaining this balance becomes not just an ethical ideal but a practical necessity for sustainable organizational integrity and human flourishing in an increasingly algorithmic world.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

References

Ajzen, I. (2020) 'The theory of planned behavior: Frequently asked questions', Human Behavior and Emerging Technologies, 2(4), pp. 314-324.

Ananny, M. and Crawford, K. (2018) 'Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability', New Media & Society, 20(3), pp. 973-989.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.F. and Rahwan, I. (2018) 'The moral machine experiment', Nature, 563(7729), pp. 59-64.

Bandura, A. (2006) 'Toward a psychology of human agency', Perspectives on Psychological Science, 1(2), pp. 164-180.

Bandura, A. (2016) Moral disengagement: How people do harm and live with themselves. Worth Publishers.

Barocas, S. and Selbst, A.D. (2016) 'Big data's disparate impact', California Law Review, 104, pp. 671-732.

Benjamin, R. (2019) Race after technology: Abolitionist tools for the New Jim Code. Polity Press.

Bijker, W.E., Hughes, T.P. and Pinch, T. (2012) The social construction of technological systems: New directions in the sociology and history of technology. MIT Press.

Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', Qualitative Research in Psychology, 3(2), pp. 77-101.

Christin, A. (2017) 'Algorithms in practice: Comparing web journalism and criminal justice', Big Data & Society, 4(2), pp. 1-14.

Coeckelbergh, M. (2020) 'Artificial intelligence, responsibility attribution, and a relational justification of explainability', Science and Engineering Ethics, 26(4), pp. 2051-2068.

Creswell, J.W. and Plano Clark, V.L. (2018) Designing and conducting mixed methods research. 3rd edn. SAGE Publications.

Cummings, M.L. (2006) 'Automation and accountability in decision support system interface design', Journal of Technology Studies, 32(1), pp. 23-31.

Darley, J.M. and Latané, B. (1968) 'Bystander intervention in emergencies: Diffusion of responsibility', Journal of Personality and Social Psychology, 8(4), pp. 377-383.

DiMaggio, P.J. and Powell, W.W. (1983) 'The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields', American Sociological Review, 48(2), pp. 147-160.

Dignum, V. (2019) Responsible artificial intelligence: How to develop and use AI in a responsible way. Springer Nature.

Dodig-Crnkovic, G. and Persson, D. (2008) 'Sharing moral responsibility with robots: A pragmatic approach', in Proceedings of the 2008 Conference on Tenth Scandinavian Conference on Artificial Intelligence: SCAI 2008. IOS Press, pp. 165-168.

Elish, M.C. (2019) 'Moral crumple zones: Cautionary tales in human-robot interaction', Engaging Science, Technology, and Society, 5, pp. 40-60.

Eubanks, V. (2018) Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. and Srikumar, M. (2020) 'Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI', Berkman Klein Center Research Publication, 2020-1.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Floridi, L. (2016) 'Faultless responsibility: On the nature and allocation of moral responsibility for distributed moral actions', Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2083), p. 20160112.

Friedman, B. and Hendry, D.G. (2019) Value sensitive design: Shaping technology with moral imagination. MIT Press.

Frith, C.D. (2014) 'Action, agency and responsibility', Neuropsychologia, 55, pp. 137-142.

Gogoll, J. and Uhl, M. (2018) 'Rage against the machine: Automation in the moral domain', Journal of Behavioral and Experimental Economics, 74, pp. 97-103.

Green, B. and Chen, Y. (2019) 'Disparate interactions: An algorithm-in-the-loop analysis of fairness in risk assessments', in Proceedings of the Conference on Fairness, Accountability, and Transparency. Association for Computing Machinery, pp. 90-99.

Ihde, D. (1990) Technology and the lifeworld: From garden to earth. Indiana University Press.

Johnson, D.G. (2006) 'Computer systems: Moral entities but not moral agents', Ethics and Information Technology, 8(4), pp. 195-204.

Kaminski, M.E. (2019) 'The right to explanation, explained', Berkeley Technology Law Journal, 34, pp. 189-218.

Klincewicz, M. (2019) 'Robotic nudges for moral improvement through Stoic practice', Techné: Research in Philosophy and Technology, 23(3), pp. 425-455.

Korsgaard, C.M. (2009) Self-constitution: Agency, identity, and integrity. Oxford University Press.

Kroll, J.A. (2018) 'The fallacy of inscrutability', Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2133), p. 20180084.

Latour, B. (2005) Reassembling the social: An introduction to actor-network theory. Oxford University Press.

Leidner, B., Castano, E., Zaiser, E. and Giner-Sorolla, R. (2010) 'Ingroup glorification, moral disengagement, and justice in the context of collective violence', Personality and Social Psychology Bulletin, 36(8), pp. 1115-1129.

Logg, J.M., Minson, J.A. and Moore, D.A. (2019) 'Algorithm appreciation: People prefer algorithmic to human judgment', Organizational Behavior and Human Decision Processes, 151, pp. 90-103.

Martin, K. (2019) 'Ethical implications and accountability of algorithms', Journal of Business Ethics, 160(4), pp. 835-850.

Matthias, A. (2004) 'The responsibility gap: Ascribing responsibility for the actions of learning automata', Ethics and Information Technology, 6(3), pp. 175-183.

Miller, T. (2019) 'Explanation in artificial intelligence: Insights from the social sciences', Artificial Intelligence, 267, pp. 1-38.

Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. and Floridi, L. (2016) 'The ethics of algorithms: Mapping the debate', Big Data & Society, 3(2), pp. 1-21.

Moore, C. and Gino, F. (2015) 'Approach, ability, aftermath: A psychological process framework of unethical behavior at work', Academy of Management Annals, 9(1), pp. 235-289.

Newman, G.E., Bullock, K. and Bloom, P. (2020) 'The psychology of delegating moral decisions to algorithms', Nature Communications, 11(5156), pp. 1-10.

Noble, S.U. (2018) Algorithms of oppression: How search engines reinforce racism. NYU Press.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Orlikowski, W.J. (2000) 'Using technology and constituting structures: A practice lens for studying technology in organizations', Organization Science, 11(4), pp. 404-428.

Pasquale, F. (2015) The black box society: The secret algorithms that control money and information. Harvard University Press.

Pinch, T.J. and Bijker, W.E. (1984) 'The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other', Social Studies of Science, 14(3), pp. 399-441.

Rest, J.R., Narvaez, D., Thoma, S.J. and Bebeau, M.J. (1999) 'DIT2: Devising and testing a revised instrument of moral judgment', Journal of Educational Psychology, 91(4), pp. 644-659.

Schlenker, B.R., Britt, T.W., Pennington, J., Murphy, R. and Doherty, K. (1994) 'The triangle model of responsibility', Psychological Review, 101(4), pp. 632-652.

Seberger, J.S. and Bowker, G.C. (2020) 'Humanistic infrastructure studies: Hyper-functionality and the experience of the absurd', Information, Communication & Society, 24(13), pp. 1-16.

Sharkey, A. (2017) 'Can robots be responsible moral agents? And why might that matter?', Connection Science, 29(3), pp. 210-216.

Star, S.L. and Griesemer, J.R. (1989) 'Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39', Social Studies of Science, 19(3), pp. 387-420.

Stevenson, M.T. (2018) 'Assessing risk assessment in action', Minnesota Law Review, 103, pp. 303-384.

Treviño, L.K., den Nieuwenboer, N.A. and Kish-Gephart, J.J. (2014) '(Un)ethical behavior in organizations', Annual Review of Psychology, 65, pp. 635-660.

Vallor, S. (2015) 'Moral deskilling and upskilling in a new machine age: Reflections on the ambiguous future of character', Philosophy & Technology, 28(1), pp. 107-124.

van de Poel, I., Royakkers, L. and Zwart, S.D. (2012) Moral responsibility and the problem of many hands. Routledge.

Veale, M., Van Kleek, M. and Binns, R. (2018) 'Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making', in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, pp. 1-14.

Verbeek, P.P. (2005) What things do: Philosophical reflections on technology, agency, and design. Pennsylvania State University Press.

Wang, D., Yang, Q., Abdul, A. and Lim, B.Y. (2020) 'Designing theory-driven user-centric explainable AI', in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, pp. 1-15.

Wong, R.Y. (2020) 'Designing for grassroots food justice: A study of three food justice organizations utilizing digital tools', Proceedings of the ACM on Human-Computer Interaction, 4(CSCW2), pp. 1-28.

Zarsky, T. (2016) 'The trouble with algorithmic decisions: An analytic road map to examine efficiency and fairness in automated and opaque decision making', Science, Technology, & Human Values, 41(1), pp. 118-132.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Appendix A: Survey Instrument and Interview Protocols

Algorithmic Moral Disengagement Survey Instrument

Introduction

Thank you for participating in this research study on how professionals interact with algorithmic decision systems. This survey will take approximately 20-25 minutes to complete. Your responses will remain confidential and will only be reported in aggregate form.

The survey asks about your experiences with algorithmic systems in your professional role. For the purpose of this survey, "algorithmic decision systems" refer to any computational tools that provide recommendations, risk scores, forecasts, or automated decisions that influence your professional decision-making.

Part I: Demographic Information

1. In which sector do you primarily work?
    - Financial Services
    - Healthcare
    - Criminal Justice
    - Other (please specify): _____
2. What is your current job title? _____
3. Gender:
    - Female
    - Male
    - Non-binary
    - Prefer not to say
    - Prefer to self-describe: _____
4. Age range:
    - 18-24
    - 25-34
    - 35-44
    - 45-54
    - 55-64
    - 65+

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Prefer not to say

5. Years of professional experience in your current field:
    - 0-5 years
    - 6-10 years
    - 11-20 years
    - 21+ years
6. How long have you been working with algorithmic decision systems in your professional role?
    - Less than 1 year
    - 1-2 years
    - 2-3 years
    - 4-5 years
    - More than 5 years

Part II: Algorithmic System Characteristics

Please answer the following questions about the primary algorithmic system you use in your professional role.

7. What type of algorithmic system do you most frequently use in your work? (Select all that apply)
    - Risk assessment tools
    - Prediction models
    - Decision recommendation systems
    - Resource allocation systems
    - Diagnostic tools
    - Other (please specify): _____
8. How would you characterize the role of this algorithmic system in your decision-making?
    - Purely advisory (provides information only)
    - Recommends actions but I make the final decision
    - Makes decisions that I can override if necessary
    - Makes automated decisions with limited override options
    - Makes fully automated decisions with no override options
9. How transparent is the algorithmic system you use?
    - Completely opaque (I don't know how it works at all)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- Somewhat opaque (I understand the basic principles but not the details)
- Moderately transparent (I understand how it works but not all the factors it considers)
- Highly transparent (I fully understand how it works and what factors it considers)

Part III: Perceived Decision-Making Autonomy Scale

Please indicate your level of agreement with the following statements about your experience with the algorithmic system.
(1 = Strongly Disagree, 7 = Strongly Agree)

10. I have significant freedom in how I use the algorithmic system's recommendations.
11. I feel constrained by the algorithmic system in my decision-making.
12. I can easily override the algorithmic system when I disagree with its recommendations.
13. My professional judgment takes precedence over the algorithmic system's recommendations.
14. I feel pressure to follow the algorithmic system's recommendations even when I disagree.
15. The algorithmic system limits my ability to use my professional expertise.
16. I have the final say in decisions that involve the algorithmic system.
17. The algorithmic system is designed to support rather than replace my judgment.

Part IV: Moral Engagement Scale

Please indicate your level of agreement with the following statements about your decision-making process when using the algorithmic system.
(1 = Strongly Disagree, 7 = Strongly Agree)

18. I carefully consider the ethical implications of decisions involving the algorithmic system.
19. I feel personally responsible for the outcomes of decisions influenced by the algorithmic system.
20. I regularly reflect on whether decisions involving the algorithmic system align with my professional values.
21. I am aware of how the algorithmic system might affect different stakeholders.
22. I actively consider alternative approaches when I have concerns about the algorithmic system's recommendations.
23. I feel engaged with the human impact of decisions influenced by the algorithmic system.
24. I consider myself morally accountable for decisions made with the algorithmic system's input.
25. I think critically about the values embedded in the algorithmic system.

Part V: Responsibility Attribution Scale

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Please indicate your level of agreement with the following statements about responsibility for decisions involving the algorithmic system.
(1 = Strongly Disagree, 7 = Strongly Agree)

26. When decisions involve the algorithmic system, responsibility is shared between me and the system.
27. If a decision based on the algorithmic system's recommendation has negative consequences, the system bears significant responsibility.
28. The developers of the algorithmic system are responsible for any unintended consequences of its recommendations.
29. My organization, rather than individual professionals, is responsible for outcomes of algorithmic decisions.
30. I am fully responsible for decisions I make, regardless of the algorithmic system's influence.
31. Following the algorithmic system's recommendations provides protection from criticism if things go wrong.
32. It would be unfair to hold me personally responsible for problematic outcomes resulting from the algorithmic system's recommendations.
33. When many people are involved in an algorithmic decision process, no one person bears full responsibility.

Part VI: Ethical Awareness Scale

Please indicate your level of agreement with the following statements about ethical aspects of using the algorithmic system.
(1 = Strongly Disagree, 7 = Strongly Agree)

34. I am aware of potential biases in the algorithmic system.
35. I consider the ethical implications of my decisions when using the algorithmic system.
36. I think about how the algorithmic system might affect vulnerable populations.
37. I reflect on whether the algorithmic system aligns with principles of fairness and justice.
38. I consider how the algorithmic system might influence power dynamics in decision processes.
39. I am mindful of situations where algorithmic recommendations might lead to discriminatory outcomes.
40. I actively consider the long-term societal implications of algorithmic decision-making in my field.
41. I try to identify ethical dilemmas created by the use of algorithmic systems.

Part VII: Algorithm Trust Scale

Please indicate your level of agreement with the following statements about your trust in the algorithmic system.
(1 = Strongly Disagree, 7 = Strongly Agree)

42. The algorithmic system generally makes good recommendations.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

43. I trust the algorithmic system more than my own judgment for certain types of decisions.
44. The algorithmic system is based on sound principles and data.
45. The algorithmic system has proven reliable over time.
46. I am skeptical about the algorithmic system's recommendations.
47. The algorithmic system sometimes makes recommendations that seem arbitrary or wrong.
48. The algorithmic system has knowledge or capabilities that complement my own expertise.
49. I would feel comfortable defending the algorithmic system's recommendations to others.

Part VIII: Organizational Context

Please indicate your level of agreement with the following statements about your organizational environment. (1 = Strongly Disagree, 7 = Strongly Agree)

50. My organization emphasizes ethical considerations in decision-making.
51. My supervisors expect me to follow the algorithmic system's recommendations.
52. My organization provides clear guidance on when to override algorithmic recommendations.
53. My colleagues regularly discuss ethical issues related to algorithmic systems.
54. My organization rewards efficiency over careful deliberation.
55. I received adequate training on the ethical use of algorithmic systems.
56. My organization has clear accountability processes for algorithmic decisions.
57. My organization values professional judgment over algorithmic recommendations.

Part IX: Implementation Approach

58. How would you characterize the implementation of the algorithmic system in your organization?
    - Top-down (leadership decided and implemented with minimal input from users)
    - Collaborative (significant input from users during design and implementation)
    - Bottom-up (users identified need and drove implementation)
    - Don't know/Wasn't involved
59. Were you consulted during the design or implementation of the algorithmic system?
    - Yes, extensively
    - Yes, somewhat
    - Minimally
    - Not at all

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

60. How would you rate your input into how the algorithmic system is used in your workflow?
    - Substantial input
    - Moderate input
    - Limited input
    - No input

Part X: Open-Ended Questions

61. Can you describe a specific situation where you felt conflicted about following the algorithmic system's recommendation? What did you do?
62. How has the algorithmic system changed how you think about your professional responsibilities?
63. What would make you more likely to take personal responsibility for decisions involving the algorithmic system?
64. Are there any other comments you would like to share about your experience with algorithmic systems in your professional role?

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Appendix B: Semi-Structured Interview Protocol

Introduction Script

Thank you for agreeing to participate in this interview. Today we'll be discussing your experiences working with algorithmic decision systems in your professional role. I'm particularly interested in understanding how these systems influence your decision-making processes and sense of professional responsibility.

The interview will take approximately 45-60 minutes. With your permission, I'll be audio-recording our conversation to ensure I accurately capture your perspectives. All your responses will remain confidential, and your identity will not be revealed in any research publications.

Before we begin, do you have any questions about the research or the interview process?

[Address any questions]

Great, let's begin.

Background Questions

1. Could you briefly describe your current professional role and responsibilities?
2. What types of algorithmic decision systems do you use in your work? How frequently do you interact with these systems?
3. How long have you been working with these algorithmic systems?
4. What training did you receive on using these systems?

Core Questions: Experiences with Algorithmic Systems

5. Could you walk me through a typical scenario where you use the algorithmic system in your decision-making process?
    - *Probe*: What information does the system provide?
    - *Probe*: How do you incorporate this information into your decision?
6. How has the algorithmic system changed your decision-making process compared to before it was implemented?
    - *Probe*: Has it changed the factors you consider?
    - *Probe*: Has it changed how much time you spend on decisions?
7. Can you describe a situation where you disagreed with the algorithmic system's recommendation?
    - *Probe*: What made you question the recommendation?
    - *Probe*: What did you do in that situation?
    - *Probe*: What factors influenced your decision to follow or override the recommendation?

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

8. How do you feel when you override the system's recommendations?
    - *Probe*: Do you experience any pressure to follow the recommendations?
    - *Probe*: How do colleagues or supervisors respond when you override recommendations?

Responsibility and Ethical Reasoning

9. When you make decisions using the algorithmic system, who do you see as responsible for the outcomes?
    - *Probe*: How is responsibility distributed between you, the system, system developers, and your organization?
    - *Probe*: Does this distribution of responsibility change depending on whether outcomes are positive or negative?
10. Has the algorithmic system changed how you think about your professional responsibilities?
    - *Probe*: Has it changed what you feel accountable for?
    - *Probe*: Has it changed how you evaluate your own performance?
11. Can you describe a situation where using the algorithmic system created an ethical dilemma for you?
    - *Probe*: How did you resolve this dilemma?
    - *Probe*: What resources or support did you draw on?
12. How do you think about the ethical implications of decisions involving the algorithmic system?
    - *Probe*: Has this changed over time as you've worked with the system?
    - *Probe*: What ethical considerations are most salient to you?

Organizational Context

13. How does your organization frame the purpose and role of the algorithmic system?
    - *Probe*: Is it presented as a decision aid or as an authority?
    - *Probe*: How is the system's accuracy or reliability discussed?
14. How does your organization handle situations where the algorithmic system makes errors or problematic recommendations?
    - *Probe*: Are there formal review processes?
    - *Probe*: How are disagreements between human judgment and algorithmic recommendations resolved?
15. How do discussions about the algorithmic system occur among your colleagues?
    - *Probe*: Do you discuss limitations or concerns about the system?
    - *Probe*: Do you share strategies for working with the system?

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

16. What organizational policies or practices influence how you use the algorithmic system?
    - *Probe*: Are there documentation requirements?
    - *Probe*: How is your use of the system evaluated?

Closing Questions

17. What changes to the algorithmic system would make it more aligned with your professional values and responsibilities?
18. What advice would you give to someone in your field who is just starting to work with similar algorithmic systems?
19. Is there anything else about your experience with algorithmic systems that you think is important for me to understand that we haven't discussed?

Closing Script

Thank you very much for sharing your experiences and insights. Your perspectives will be valuable for understanding how algorithmic systems influence professional decision-making and ethical reasoning.

Do you have any questions for me about the research or how your interview will be used?

[Address any questions]

If you think of anything else you'd like to add or if you have any questions later, please feel free to contact me.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Appendix C: Experimental Vignette Protocol

Introduction

Thank you for participating in this research study. In this session, you will read several scenarios involving algorithmic decision systems similar to those used in your professional field. After each scenario, you will be asked to respond to questions about how you would think, feel, and act in the described situation.

There are no right or wrong answers. We are interested in your professional judgment and reasoning about these scenarios. Your responses will remain confidential and will only be used for research purposes.

Instructions

You will read 6 scenarios. Each scenario describes a situation involving an algorithmic decision system in your professional field. The scenarios vary in:

- The degree of constraint imposed by the algorithmic system
- The transparency of the algorithmic system
- The outcome of the decision

After each scenario, please answer questions about responsibility attribution, ethical evaluation, and your satisfaction with the decision process.

Sample Vignettes for Financial Services Professionals

Vignette 1: High Constraint, Opaque Algorithm, Negative Outcome

You are reviewing a loan application from a small business owner seeking to expand their successful local restaurant. The business has shown consistent growth over the past three years and has a good repayment history on a previous smaller loan.

When you enter the application details into the algorithmic loan evaluation system, it generates a "high risk" score of 720 (above the 700 threshold). According to organizational policy, applications scoring above 700 must be declined unless approved by a senior manager. The system does not provide detailed reasoning for its risk assessment beyond listing "business sector volatility" as a primary factor.

Following protocol, you decline the loan. Three months later, you learn that the business secured financing from a competitor bank and has successfully expanded, increasing revenue by 40% and hiring five new employees.

Questions:

1. To what extent do you feel personally responsible for the decision to decline the loan? (1-7 scale)
2. To what extent is each of the following responsible for the decision outcome:
    - You
    - The algorithmic system

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

- The system developers
- Your organization's policies
- The senior management

3. How ethically problematic do you find this situation? (1-7 scale)
4. How satisfied are you with the decision process in this scenario? (1-7 scale)
5. What would you have done differently in this situation, if anything?

Vignette 2: Low Constraint, Transparent Algorithm, Positive Outcome

You are evaluating a mortgage application from a couple with moderate income but excellent credit history. When you enter their information into the mortgage evaluation system, it calculates a "moderate risk" score of 620 (below the 650 high-risk threshold).

The system displays a detailed breakdown of its assessment, showing that while the applicants' debt-to-income ratio (45%) is slightly higher than ideal, their excellent payment history, employment stability, and moderate loan-to-value ratio significantly mitigate this risk. The system recommends approval with slightly higher interest rates to offset the risk.

Your organization's policy allows you full discretion for applications scoring below 650. Based on your review of their complete financial situation and the system's transparent analysis, you approve the mortgage at the standard interest rate. Over the next two years, the applicants make all payments on time and refer two friends to your bank.

Questions:

1. To what extent do you feel personally responsible for the decision to approve the mortgage? (1-7 scale)
2. To what extent is each of the following responsible for the decision outcome:
    - You
    - The algorithmic system
    - The system developers
    - Your organization's policies
    - The senior management
3. How ethically appropriate do you find this situation? (1-7 scale)
4. How satisfied are you with the decision process in this scenario? (1-7 scale)
5. What factors most influenced your sense of responsibility in this scenario?

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Sample Vignettes for Healthcare Professionals

Vignette 3: High Constraint, Transparent Algorithm, Positive Outcome

You are treating a 58-year-old patient with symptoms suggesting a possible sepsis infection. You enter the patient's vital signs and test results into the hospital's sepsis detection algorithm, which calculates a 78% probability of sepsis (above the 70% protocol threshold).

The system displays a detailed explanation of its assessment, showing how specific combinations of the patient's elevated heart rate, respiration rate, decreased blood pressure, and laboratory values contribute to the high probability score. According to hospital protocol, scores above 70% require immediate initiation of the sepsis treatment bundle.

Although you initially thought the patient might have a less severe condition, you follow the protocol and initiate the treatment bundle. Within 12 hours, the patient's condition improves significantly, and later test results confirm the sepsis diagnosis. The early intervention likely prevented serious complications.

Questions:

1. To what extent do you feel personally responsible for the positive patient outcome? (1-7 scale)
2. To what extent is each of the following responsible for the decision outcome:
    - You
    - The algorithmic system
    - The system developers
    - Your hospital's protocols
    - The medical team
3. How ethically appropriate do you find this situation? (1-7 scale)
4. How satisfied are you with the decision process in this scenario? (1-7 scale)
5. How would you feel about using this system for future patients?

Vignette 4: Low Constraint, Opaque Algorithm, Negative Outcome

You are evaluating a 45-year-old patient with atypical chest pain. After initial tests, you enter the patient's information into the cardiac risk assessment algorithm, which calculates a "low risk" score of 22% (below the 30% threshold for recommended additional testing).

The system does not provide detailed reasoning for its assessment beyond listing "multiple factors" as the basis for the score. Hospital guidelines suggest that physicians use their judgment for patients scoring below 30%.

Based on your clinical experience and the algorithmic assessment, you decide against ordering additional cardiac tests and prescribe anti-inflammatory medication for presumed costochondritis (chest wall inflammation). Three days later, the patient returns with a myocardial infarction (heart attack) that could have been detected with additional testing.

Questions:

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

1. To what extent do you feel personally responsible for the negative patient outcome? (1-7 scale)
2. To what extent is each of the following responsible for the decision outcome:
    - You
    - The algorithmic system
    - The system developers
    - Your hospital's guidelines
    - The medical team
3. How ethically problematic do you find this situation? (1-7 scale)
4. How satisfied are you with the decision process in this scenario? (1-7 scale)
5. What would you have done differently in this situation, if anything?

Sample Vignettes for Criminal Justice Professionals

Vignette 5: High Constraint, Opaque Algorithm, Positive Outcome

You are a judge determining sentencing for a defendant convicted of non-violent drug possession. The defendant has one prior conviction for a similar offense five years ago. When you enter the case information into the recidivism risk assessment algorithm, it generates a "high risk" score of 8.2 on a 10-point scale.

The algorithm does not provide detailed reasoning for its assessment beyond indicating "criminal history patterns" as significant factors. Court guidelines strongly recommend incarceration for defendants scoring above 7.5, with limited judicial discretion.

Following the guidelines, you sentence the defendant to 18 months incarceration with mandatory drug treatment. During this period, the defendant completes treatment successfully, earns educational credentials, and upon release, maintains sobriety and employment for the next three years without reoffending.

Questions:

1. To what extent do you feel personally responsible for the sentencing decision? (1-7 scale)
2. To what extent is each of the following responsible for the decision outcome:
    - You
    - The algorithmic system
    - The system developers
    - The court guidelines
    - The corrections system
3. How ethically appropriate do you find this situation? (1-7 scale)
4. How satisfied are you with the decision process in this scenario? (1-7 scale)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

5. What factors most influenced your sense of responsibility in this scenario?

Vignette 6: Low Constraint, Transparent Algorithm, Negative Outcome

You are determining bail conditions for a defendant charged with burglary. The defendant has stable employment, family ties to the community, and no prior felony convictions. When you enter the case information into the pretrial risk assessment algorithm, it calculates a "moderate flight risk" score of 42%.

The system provides a detailed breakdown showing how it weighed factors including the defendant's recent change of address (negative), steady employment (positive), lack of prior failures to appear (positive), and nature of the current charge (negative). Court policy allows full judicial discretion for defendants with moderate risk scores (30-60%).

After reviewing all information and the algorithm's transparent analysis, you set bail at 5,000,lowerthantheprosecutor'srequestof5,000, lower than the prosecutor's request of 5,000,lowerthantheprosecutor'srequestof15,000. The defendant posts bail but fails to appear for trial and is arrested in another state six weeks later.

Questions:

1. To what extent do you feel personally responsible for the defendant's failure to appear? (1-7 scale)
2. To what extent is each of the following responsible for the decision outcome:
    - You
    - The algorithmic system
    - The system developers
    - The court policies
    - The defendant
3. How ethically problematic do you find this situation? (1-7 scale)
4. How satisfied are you with the decision process in this scenario? (1-7 scale)
5. What would you have done differently in this situation, if anything?

Debriefing

Thank you for completing this study. The scenarios you evaluated were designed to examine how different characteristics of algorithmic systems influence perceptions of responsibility and ethical reasoning.

Specifically, we are investigating:

1. How the level of constraint imposed by algorithmic systems affects professionals' sense of responsibility
2. How algorithmic transparency influences ethical evaluation and decision satisfaction
3. How outcome valence (positive vs. negative) impacts responsibility attribution

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Your responses will help us understand these dynamics and develop guidelines for the ethical implementation of algorithmic systems in professional contexts.

Do you have any questions about the study or your participation?

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Generational Divides and Gender Dynamics: Decoding the Multifaceted Drivers of Employee Engagement

Jonathan H. Westover [a b c *]

[a] Utah Valley University, UT, USA
[b] Future State University, CA, USA
[c] Nexus Institute for Work & AI, Catalyst Center for Work Innovation, UT, USA
[*] Correspondence: jon.westover@gmail.com

Received April 15, 2026
Accepted for publication July 22, 2026
Published Early Access August 5, 2026

doi.org/10.70175/futureofworkjournal.2026.1.1.4

Abstract: *This study aims to provide a more nuanced, age- and gender-specific understanding of the key drivers of employee engagement. By analyzing survey data from over 500 U.S. employees, the research evaluates the relative influence of traditional predictors like basic needs fulfillment alongside evolving constructs like "worker activation" - discretionary commitments nurtured through empowering organizational cultures. The findings reveal significant generational and gender differences in employee engagement levels and the salience of various engagement determinants. Baby Boomers and men exhibit higher overall engagement as well as stronger senses of purpose, belonging, leadership efficacy, and organizational commitment compared to younger generations and women. Basic needs and teamwork factors are more influential for younger cohorts and women, while individual determinants and growth opportunities matter more for older generations and men. These insights can inform the development of tailored, workforce-responsive strategies to foster commitment, performance, and well-being across an organization's multigenerational, gender-diverse employees, as understanding the distinct engagement drivers for different generational and gender groups is critical for optimizing discretionary effort and organizational success.*

Keywords: employee engagement, generational differences, gender differences, workplace diversity, organizational commitment, worker activation, workforce management, engagement drivers

Suggested Citation:

Westover, Jonathan H. (2026). Generational Divides and Gender Dynamics: Decoding the Multifaceted Drivers of Employee Engagement. *Future of Work: The Journal of Labor Transformation, Technology Integration, and Human Adaptation*, 1(1). doi.org/10.70175/futureofworkjournal.2026.1.1.4

Employee engagement has emerged as a critical factor influencing organizational success, with heightened engagement predicting lower turnover, greater productivity, and improved safety outcomes. As such, understanding the drivers of discretionary effort in the workplace remains a top priority for leaders and HR professionals. However, employee experiences and the dynamics shaping engagement are likely to differ based on various personal attributes, including generational cohort and gender.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Existing research has yielded mixed and sometimes contradictory findings on the relationships between generation, gender, and employee engagement. While some studies have identified distinct engagement predictors across age groups, the nuances of how factors like basic needs fulfillment, individual determinants, teamwork dynamics, and growth opportunities impact discretionary effort for different generations and genders remain under-explored. For example, research suggests that Millennial women may place greater emphasis on role clarity and supervisor support, while Baby Boomer men appear to be more motivated by the meaningfulness of their work. However, the relative salience of these engagement drivers has not been comprehensively analyzed across all generational cohorts and gender groups.

To address these gaps, the current study aims to provide a more nuanced, age- and gender-specific understanding of the key drivers of employee engagement. By analyzing survey data from over 500 U.S. employees, this research evaluates the relative influence of traditional predictors like basic needs fulfillment alongside evolving constructs like "worker activation" - discretionary commitments nurtured through empowering organizational cultures (Westover & Andrade, 2024). Importantly, the study examines these engagement dynamics separately for each generational cohort (Baby Boomers, Generation X, Millennials, and Generation Z) and by gender, providing critical insights into the unique predictors of engagement across age and gender groups.

Overall, the findings aim to provide strategic direction for organizations seeking to optimize engagement and discretionary effort within their multigenerational, gender-diverse workforces. Understanding the distinct engagement drivers for different generational and gender groups can inform the development of tailored, workforce-responsive strategies to foster commitment, performance, and well-being.

LITERATURE REVIEW

Overview of Employee Engagement

Employee engagement refers to the level of an employee's emotional commitment, enthusiasm, and passion for their work. Highly engaged employees exhibit a strong sense of purpose, a willingness to go "above and beyond" in their roles, and a deep connection to the organization's mission and values (Schaufeli & Bakker, 2004). In contrast, disengaged workers lack motivation, show reduced productivity, and have a higher propensity for turnover (Macey & Schneider, 2008).

Employee engagement is critical for organizational success, as engaged employees are emotionally committed to their work and the organization, demonstrating high levels of enthusiasm, dedication, and discretionary effort (Gallup, 2022; Kahn, 1990). Disengaged employees, on the other hand, are physically present but psychologically absent, contributing minimally to organizational goals (Kahn, 1990; Macey & Schneider, 2008).

Scholars have conceptualized employee engagement as a multifaceted construct involving cognitive, emotional, and behavioral components (Kahn, 1990; Schaufeli et al., 2002; Shuck & Wollard, 2010). Kahn (1990) defined engagement as the "harnessing of organization members' selves to their work roles," while Schaufeli et al. (2002) described it as a "positive, fulfilling, work-related state of mind" characterized by vigor, dedication, and absorption.

Several theoretical frameworks have been used to explain the drivers and processes underlying employee engagement, including the job demands-resources (JD-R) model (Demerouti et al., 2001; Schaufeli & Bakker, 2004), the conservation of resources (COR) theory (Hobfoll, 1989), the social exchange theory (SET) (Blau, 1964), and the self-determination theory (SDT) (Ryan & Deci, 2000).

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Researchers have identified a range of individual, job-related, and organizational factors that contribute to employee engagement, such as personality traits, job characteristics, organizational support, and work-life balance (Bakker et al., 2012; Christian et al., 2011; Saks, 2006). Engagement has been linked to positive individual and organizational outcomes, including higher job satisfaction, performance, and financial success (Halbesleben, 2010; Harter et al., 2002).

Additionally, various instruments, such as the Utrecht Work Engagement Scale (UWES), the Gallup Q12, and the Intellectual, Social, and Affective Engagement Scale (ISA), have been developed to measure and assess employee engagement (Schaufeli et al., 2002; Rich et al., 2010). The measurement and assessment of engagement can be influenced by contextual factors, requiring careful consideration of the appropriate tools for a given organizational context (Albrecht, 2010; Schaufeli & Bakker, 2004).

Key Determinants of Employee Engagement

Employee engagement has emerged as a critical factor for organizational success, with extensive research highlighting its impact on various outcomes (Kahn, 1990; Saks, 2006). This literature review provides an overview of the key determinants of employee engagement, drawing from existing theoretical frameworks and empirical studies.

At the foundational level, employee engagement is influenced by the fulfillment of basic needs, as outlined in Maslow's Hierarchy of Needs (Maslow, 1943). Employees require their physiological, safety, and belongingness needs to be met in order to be engaged (Kahn, 1990; May et al., 2004).

Beyond basic needs, individual factors also play a significant role in shaping engagement. These include personality traits, psychological states, and personal resources (Bakker et al., 2014; Christian et al., 2011; Xanthopoulou et al., 2007).

The work environment, particularly the quality of interpersonal relationships and team dynamics, can also influence engagement. Factors such as social support, team cohesion, and collaborative climate are associated with higher engagement levels (Breevaart et al., 2014; Kahn, 1990; Saks, 2006).

Opportunities for growth and development are crucial in fostering employee engagement. Employees are more engaged when their jobs provide autonomy, skill variety, task significance, and feedback, as well as opportunities for learning and perceived meaningfulness (Bakker & Demerouti, 2008; Hackman & Oldham, 1976; Kahn, 1990).

Finally, the degree to which employees are "activated" or empowered in their work roles can also impact engagement. Factors such as autonomy, role clarity, and feedback and recognition are positively associated with engagement (Christian et al., 2011; Kahn, 1990; Saks, 2006).

By understanding these multifaceted determinants of employee engagement, organizations can develop strategies to foster a highly engaged workforce, leading to improved organizational performance and outcomes.

Generational Differences in Employee Engagement

Over the past two decades, employee engagement has become an increasingly important topic in organizational research and practice (Gallup, 2022; Kahn, 1990). Engaged employees are emotionally committed to their work and organization, and demonstrate high levels of enthusiasm, dedication, and discretionary effort. Engagement has been linked to a variety of positive individual and organizational

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

outcomes, including increased productivity, customer satisfaction, and financial performance (Harter et al., 2002; Saks, 2006).

As the workforce becomes more diverse, with multiple generations working side-by-side, there is growing interest in understanding how generational differences may influence employee engagement (Becton et al., 2014; Costanza & Finkelstein, 2015). Generational cohorts are typically defined as groups of individuals born within a similar time period, who share common experiences, values, and attitudes shaped by the social, economic, and political events of their formative years (Mannheim, 1952; Parry & Urwin, 2011). The most commonly recognized generational cohorts in the current workforce include *Baby Boomers* (born 1946-1964), *Generation X* (born 1965-1980), *Millennials* (born 1981-1996), and *Generation Z* (born 1997-2012) (Gallup, 2022; Pew Research Center, 2019).

Several theoretical frameworks have been proposed to explain the origins and dynamics of generational differences in the workplace. The generational theory, developed by Strauss and Howe (1991), posits that generations are shaped by a cyclical pattern of social, economic, and political events. The life-stage theory suggests that differences are primarily driven by the stage of life and career development, rather than unique historical experiences (Schaie, 1965; Super, 1957). The socialization theory emphasizes the role of cultural and social influences in shaping generational differences (Mannheim, 1952; Parry & Urwin, 2011).

The existing literature on generational differences in employee engagement has produced mixed and sometimes contradictory findings (Becton et al., 2014; Costanza & Finkelstein, 2015). Researchers have identified several factors that may contribute to these differences, including work values, career expectations, management preferences, and work-life balance needs (Costanza & Finkelstein, 2015; Twenge et al., 2010). Organizational factors, such as culture, HR practices, and leadership styles, can also play a significant role in shaping generational differences in employee engagement (Costanza & Finkelstein, 2015).

Gender Differences in Employee Engagement

Recent research paints a nuanced picture of gender differences in employee engagement. A global survey found that women generally report higher levels of engagement than men, exhibiting greater commitment, enthusiasm, and positive impact within their organizations (Frumar & Truscott-Smith, 2024). This pattern holds true for women in project management, managerial, and individual contributor roles.

However, the gender engagement gap reverses at the senior leadership level, where women are less engaged than their male counterparts. This may be attributed to factors like isolation, lack of emotional support, and the absence of close relationships that tend to characterize senior-level positions, leading women to stay in these roles for shorter durations (Frumar & Truscott-Smith, 2024).

Occupational self-efficacy also plays a role, with studies showing that men have higher career aspirations even when their self-efficacy is on par with women's (Hartman & Barber, 2020). Women, particularly those with low or moderate self-efficacy, may benefit from coaching, mentoring, and career development support to overcome this barrier.

The relationship between gender and engagement is also context-dependent. While some studies found no differences between men and women in certain work settings, others reported men experiencing higher engagement, commitment, and inclusion in roles like IT in India (Mulaudzi & Takawira, 2015; Sharma et al., 2017; Nobles, 2023; Zoe Talent Solutions, 2024).

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

To address these disparities, organizations can implement strategies like flexible work arrangements, leadership opportunities, mentorship, networking groups, and inclusive language training to support women's engagement and advancement, particularly in senior roles (Frumar & Truscott-Smith, 2024). Fostering a diverse organizational climate and strong team relationships can also help counter the isolation that women may face at the top.

HYPOTHESES

The existing literature on generational and gender differences in employee engagement has produced mixed findings, with limited understanding of how basic needs, individual determinants, teamwork factors, and growth aspects impact engagement for different age cohorts and between men and women. Research suggests younger generations and women may place greater emphasis on factors like role clarity, supervisor support, and teamwork, while older cohorts and men may be more motivated by the meaningfulness of their work and individual growth opportunities. However, the relative salience of these engagement predictors across generations and genders remains underexplored. Investigating the combined effects of generational and gender differences on the drivers of employee engagement can inform the development of tailored, responsive strategies to foster commitment, performance, and well-being across an organization's diverse workforce.

To address these gaps in the literature, we propose the following hypotheses to examine potential variations in the drivers of employee engagement across generational cohorts and by gender:

- **Hypothesis 1**: Employee engagement levels will differ significantly between generational cohorts and gender groups.
- **Hypothesis 2a**: The influence of basic needs and individual contribution variables on employee engagement will vary across generational cohorts and gender.
- **Hypothesis 2b**: Basic needs determinants will be more salient in predicting engagement for younger generations (e.g., Millennials, Gen Z) compared to older cohorts, and for women compared to men.
- **Hypothesis 2c**: Individual determinants will be more salient in predicting engagement for older generations (e.g., Baby Boomers, Gen X) compared to younger cohorts, and for men compared to women.
- **Hypothesis 3**: Teamwork determinants will be more salient in predicting employee engagement for younger generations (e.g., Millennials, Gen Z) than older cohorts, and for women compared to men.
- **Hypothesis 4**: Growth determinants will be more salient in predicting employee engagement for older generations (e.g., Baby Boomers, Gen X) than younger cohorts, and for men compared to women.
- **Hypothesis 5**: Worker activation determinants will be more salient in predicting employee engagement for younger generations (e.g., Millennials, Gen Z) than older cohorts, and for women compared to men.

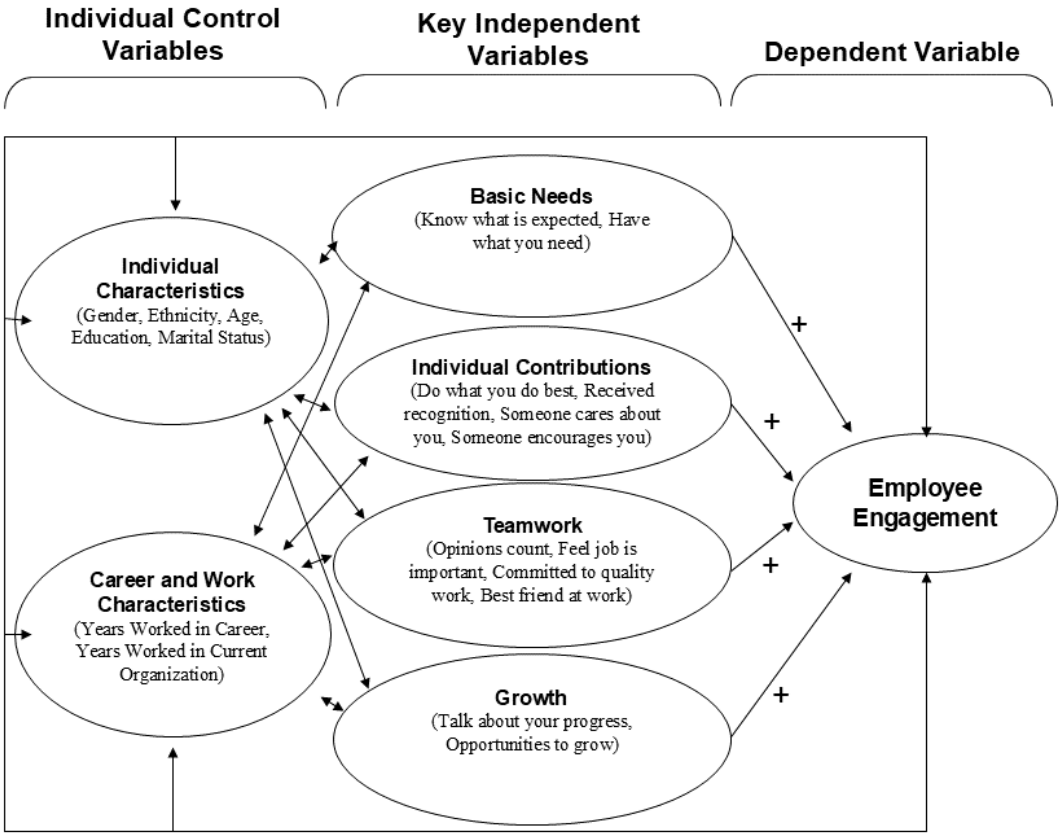
By testing these hypotheses, the study aims to provide a more nuanced, age- and gender-specific understanding of the key drivers of employee engagement. This knowledge can inform the development of tailored, generationally- and gender-responsive strategies to foster commitment, performance, and well-being across an organization's diverse workforce.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

RESEARCH MODEL AND DESIGN

Drawing on prominent employee engagement frameworks, we developed a comprehensive online survey to investigate the evolving dynamics of the modern workplace. The questionnaire explored a range of factors, including basic needs, individual contributions, teamwork dynamics, growth opportunities, and employee activation. Using a stratified random sampling approach across the United States, we collected a final dataset of 566 employee responses, providing a reliable and generalizable snapshot of the contemporary work experience. By blending insights from seminal and more recent engagement research, the survey design allowed us to obtain a holistic understanding of the multifaceted influences on employee attitudes and behaviors in today's rapidly changing work environment.

FIGURE 1: RESEARCH MODEL



Operationalization of Variables

We operationalized the study variables according to the approach of Westover and Andrade (2024) and added new survey questions, which allowed us to introduce additional variables in the analysis. See Table 1 below.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

TABLE 1: STUDY VARIABLES AND MEASUREMENTS

Variable	Item
Dependent Variable	
Employee engagement	"Overall, how engaged are you in your (main) job?" (1) not at all engaged to (10) extremely engaged
Worker Engagement	
Know what is expected	"Do you know what is expected of you at work?" (1) strongly disagree to (5) strongly agree
Have what you need	"Do you have the materials and equipment to do your work right?" (1) strongly disagree to (5) strongly agree
Do what you do best	"I Have the opportunity to do what I do best every day." (1) strongly disagree to (5) strongly agree
Received recognition	"In the last seven days, have you received recognition or praise for doing good work?" (1) strongly disagree to (5) strongly agree
Someone cares about you	"Does your supervisor, or someone at work, seem to care about you as a person?" (1) strongly disagree to (5) strongly agree
Someone encourages you	"Is there someone at work who encourages your development?" (1) strongly disagree to (5) strongly agree
Opinions count	"At work, do your opinions seem to count?" (1) strongly disagree to (5) strongly agree
Feel job is important	"Does the mission/purpose of your company make you feel your job is important?" (1) strongly disagree to (5) strongly agree
Committed to quality work	"Are your associates (fellow employees) committed to doing quality work?" (1) strongly disagree to (5) strongly agree
Best friend at work	"Do you have a best friend at work?" (1) strongly disagree to (5) strongly agree
Talk about your progress	"In the last six months, has someone at work talked to you about your progress?" (1) strongly disagree to (5) strongly agree
Opportunities to grow	"In the last year, have you had opportunities to learn and grow?" (1) strongly disagree to (5) strongly agree
Understanding of Meaning and Purpose	
Meaningful work	"I have a good sense of what makes my job meaningful." (1) strongly disagree to (5) strongly agree
Purposeful work	"I have discovered work that has a satisfying purpose." (1) strongly disagree to (5) strongly agree
Sense of Belonging	"I believe that my work group is where I am meant to be." (1) strongly disagree to (7) strongly agree
Leadership Efficacy	"I see myself as a leader." (1) strongly disagree to (5) strongly agree

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Organizational Commitment	"I would be very happy to spend the rest of my career with this organization." (1) strongly disagree to (5) strongly agree
Controls	Dummy variables for race, ethnicity, education level, marital status, and state of residence; Continuous variables for birth year, full-time years worked in career, and years worked in current organization.

Statistical Methodology

The analysis employed a multi-stage approach to examine the relationship between employee engagement and various workplace factors across generational cohorts and gender. First, we conducted preliminary and descriptive analyses of worker engagement and activation variables for the full sample, as well as by generational group (e.g., Baby Boomers, Generation X, Millennials, Generation Z) and gender.

This initial step provided insights into potential differences in engagement and activation levels between the age cohorts and between men and women. Next, we tested for statistically significant differences in employee engagement across the generational groups and between genders using t-test analyses (Hypothesis 1). This allowed us to determine if engagement levels differed significantly between the various age cohorts and between men and women.

Building on these initial findings, we then examined generational- and gender-specific OLS and ordered probit regression models to evaluate the relative contribution of employee basic needs, individual contributions, teamwork, and growth factors on engagement for each age and gender group (Hypotheses 2a-2c, 3, 4). This analysis shed light on how the salience of these engagement determinants varied across the different generational cohorts and between men and women.

Finally, we tested for statistically significant differences between the generational groups and between genders in the impact of worker activation dimensions (purposeful work, sense of belonging, leadership efficacy, organizational commitment) on employee engagement using moderation analyses (Hypotheses 5a-5d). This allowed us to identify any generational and gender nuances in how these activation-related factors shaped discretionary effort.

By employing this multi-stage analytical approach, we were able to provide a comprehensive, age- and gender-specific understanding of the key drivers of employee engagement. The combination of descriptive, comparative, and regression-based techniques enabled us to uncover both the overarching patterns as well as the nuanced, generational and gender variations in what inspires discretionary effort across the workforce.

RESULTS

Participant Demographics

The study's sample consisted of 566 respondents with diverse demographic characteristics. The majority of participants identified as White (67.67%), followed by Black or African American (19.96%) and Asian (9.72%), with smaller proportions identifying as Native American/Alaska Native (0.35%), Native Hawaiian/Pacific Islander (0.71%), and Other (1.59%). In terms of education level, the largest group had a bachelor's degree (34.10%), followed by some college but no degree (26.11%), a high school diploma (17.05%), a master's degree (17.23%), and a doctoral degree (4.44%). Only 1.07% had less than a high school education. Regarding ethnicity, the majority of respondents (88.34%) indicated they were not of Hispanic, Latino, or

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Spanish origin, while 11.66% did identify as such. For marital status, most participants were married or cohabitating (62.70%), with 36.59% reporting they were single, and a small percentage (0.71%) preferring not to disclose their marital status. The gender breakdown of the sample was 53.89% female and 46.11% male. In terms of generation, the largest group was Gen X (34.10%), followed by Millennials (33.00%), Baby Boomers (23.00%), and Gen Z (9.80%). The mean birth year of respondents was 1977.34 (SD = 13.99), and on average they had worked 20.57 years (SD = 13.92) in their career and 13.94 years (SD = 86.29) in their current organization. Overall, the sample represented a diverse set of participants in terms of race, education, ethnicity, marital status, gender, and generation, providing a robust foundation for the study's findings.

TABLE 2: PARTICPANT DEMOGRAPHICS

Race of Respondent	Freq.	Percent
White	383	67.67
Black or African American	113	19.96
Asian	55	9.72
Native American or Alaska Native	2	0.35
Native Hawaiian or Pacific Islander	4	0.71
Other	9	1.59
Total	566	100
Education Level of Respondent	Freq.	Percent
Less than high school	6	1.07
High school diploma	96	17.05
Some college, but no degree	147	26.11
Bachelor's degree	192	34.1
Master's degree	97	17.23
Doctoral degree	25	4.44
Total	563	100
Ethnicity of Respondent	Freq.	Percent
Hispanic or Latino or Spanish Origin	66	11.66
Not Hispanic or Latino or Spanish Origin	500	88.34
Total	566	100
Marital Status of Respondent	Freq.	Percent
Married or cohabitating	353	62.7

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Single	206	36.59
Prefer not to say	4	0.71
Total	563	100
Gender of Respondent	Freq.	Percent
Female	305	53.89
Male	261	46.11
Total	566	100
Generation of Respondent	Freq.	Percent
Baby Boomer	129	23.0%
Gen X	191	34.1%
Millennial	185	33.0%
Gen Z	55	9.8%
Total	560	100
Other Demographic Variables	Mean	Std. Dev.
Birth year	1977.34	13.99
Full-time years worked in career	20.57	13.92
Years worked in current organization	13.94	86.29

Descriptive Results

The data presented in Table 3 reveals significant generational and gender differences in employee engagement and activation within the organization. Examining overall engagement levels, Baby Boomers stand out as the most engaged, with average scores of 8.44 for women and 8.49 for men. This is followed by Gen X (7.85 for women, 8.35 for men), Millennials (7.44 for women, 7.82 for men), and Gen Z (7.31 for women, 7.75 for men). These differences in engagement across the generational cohorts are statistically significant, indicating that employee engagement declines as we move from older to younger generations.

Looking within each generation, there is also a clear gender gap, with men exhibiting higher levels of engagement than their female counterparts. This gender disparity is statistically significant, with men scoring 8.19 on average compared to 7.70 for women overall.

Delving deeper into the subdimensions of engagement, the data shows no notable generational or gender differences in the more basic elements, such as role clarity, access to necessary resources, and opportunities to leverage one's strengths. However, distinct patterns emerge in areas like recognition, supervisor support, having one's opinions valued, and perceiving the organization's mission as important. In these facets,

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

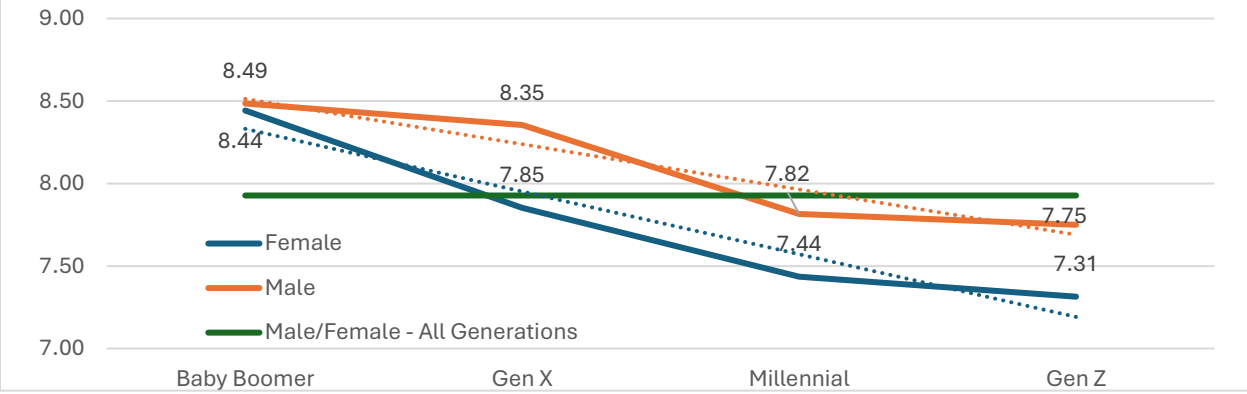
Baby Boomers and men consistently report more positive experiences than younger generations and women, respectively.

A similar trend is observed in the employee activation measures. Baby Boomers and men exhibit a stronger sense of purpose, belonging, leadership efficacy, and organizational commitment compared to other generational cohorts and women. These gaps in worker activation are also statistically significant.

TABLE 3: VARIABLE MEANS AND TEST OF DIFFERENCES, BY GENERATION & GENDER

	Baby Boomer		Gen X		Millennial		Gen Z		All		T Statistic & p-value for sig. diff		
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	*t*	p-value	df
Dependent Variable													
Employee Engagement	8.44	8.49	7.85	8.35	7.44	7.82	7.31	7.75	7.70	8.19	-2.861**	0.002	563
Employee Engagement Questions													
Do you know what is expected of you at work?	4.75	4.76	4.69	4.61	4.62	4.51	4.54	4.25	4.64	4.60	*n.s.*	*n.s.*	*n.s.*
Do you have the materials and equipment to do your work right?	4.54	4.54	4.37	4.25	4.44	4.33	4.34	3.60	4.40	4.30	*n.s.*	*n.s.*	*n.s.*
At work, do you have the opportunity to do what you do best every day?	4.31	4.44	4.17	4.26	4.15	4.25	4.09	3.90	4.16	4.28	*n.s.*	*n.s.*	*n.s.*
In the last seven days, have you received recognition or praise for doing good work?	3.43	3.68	3.46	3.71	3.57	3.99	3.37	3.40	3.47	3.75	-2.399**	0.008	563
Does your supervisor, or someone at work, seem to care about you as a person?	4.13	4.29	4.02	4.19	4.17	4.26	3.71	3.90	4.04	4.22	-2.035*	0.021	563
Is there someone at work who encourages your development?	3.80	3.90	3.74	3.88	3.99	4.03	3.86	4.05	3.84	3.94	*n.s.*	*n.s.*	*n.s.*
At work, do your opinions seem to count?	3.90	4.18	3.55	4.06	3.97	4.04	3.94	4.15	3.80	4.10	-3.033***	0.001	563
Does the mission/purpose of your company make you feel your job is important?	3.98	4.21	3.91	4.13	3.92	4.01	3.77	4.00	3.89	4.11	-2.374**	0.009	563
Are your associates (fellow employees) committed to doing quality work?	4.02	4.18	3.80	4.20	4.00	4.17	3.80	4.40	3.90	4.20	-3.695***	0.000	563
Do you have a best friend at work?	3.21	3.37	3.18	3.47	3.42	3.59	3.49	3.65	3.30	3.49	-1.581*	0.05	563
In the last six months, has someone at work talked to you about your progress?	3.25	3.43	3.37	3.72	3.84	3.87	3.60	4.20	3.53	3.72	-1.685*	0.046	563
In the last year, have you had opportunities to learn and grow?	3.62	3.79	3.66	4.13	4.05	4.08	3.60	4.05	3.77	4.02	-2.462**	0.007	563
Employee Activation Questions													
I have a good sense of what makes my job meaningful.	4.31	4.31	4.00	4.03	4.01	3.92	3.74	3.85	4.02	4.06	*n.s.*	*n.s.*	*n.s.*
I have discovered work that has a satisfying purpose.	4.11	4.13	3.91	4.07	3.77	3.92	3.74	3.95	3.87	4.04	-1.850*	0.032	563
I believe that my work group is where I am meant to be.	5.00	5.59	4.88	5.29	5.12	5.17	4.40	5.05	4.91	5.32	-2.860**	0.002	562
I see myself as a leader.	3.36	3.94	3.97	4.06	4.06	4.12	3.74	4.45	3.85	4.08	-1.781*	0.038	563
I would be very happy to spend the rest of my career with this organization.	5.38	5.59	4.89	5.30	4.94	5.14	4.46	4.85	4.93	5.30	-2.404**	0.008	562

FIGURE 2: MEAN EMPLOYEE ENGAGEMENT, BY GENERATION & GENDER



The findings from this analysis suggest that organizations should consider implementing tailored engagement strategies to address the diverse needs and perspectives of their multigenerational and gender-diverse workforce. Adopting a more nuanced, demographic-specific approach may be key to enhancing engagement and retention across the organization.

Regression Results

Building on the work of Westover and Andrade (2024), we conducted a series of regression analyses to examine the relationship between employee engagement and a range of independent variables. The initial model (Table 4) explored the impact of factors like basic needs, individual contributions, teamwork, growth opportunities, and control variables on engagement, analyzing the results by gender and geographic location. In the second model (Table 5), we expanded the analysis to consider the combined effect of all control and independent variables on engagement. Importantly, we also incorporated a set of "employee activation"

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

variables for each gender, location, and the overall sample. The inclusion of these activation factors led to many variables from the first model becoming non-significant.

As a result, the final model (Table 6) focuses solely on the most impactful engagement and activation variables, which we believe represents the optimal predictive framework. This parsimonious yet comprehensive model allows us to identify the key drivers of employee engagement across the organization. By taking this stepwise approach and progressively refining the regression analyses, we were able to isolate the most salient factors influencing engagement within this workforce. This data-driven process enabled us to develop what we consider to be the "best" model for understanding and promoting employee engagement in the context of this study.

Model 1: Original Employee Engagement Model

In Table 4, the results offer valuable insights into the antecedents of engagement for each generation, by gender. The regression analysis confirms the significant generational and gender disparities in employee engagement observed in the previous analysis. Among Baby Boomers, being female is associated with higher engagement on several subdimensions, including knowing role expectations (β=0.242, $p<0.05$), having the necessary materials (β=-0.222, $p<0.05$), and having a best friend at work (β=0.193, $p<0.05$). In contrast, for Generation X, being female is negatively related to having the necessary materials (β=-0.160, $p<0.05$) and receiving progress feedback (β=-0.181, $p<0.05$). For Millennials, being female is positively linked to supervisor support (β=0.222, $p<0.05$) and opportunities for learning and growth (β=0.235, $p<0.05$). Across generations, men generally report higher engagement in areas like role clarity, mission/purpose, and having a best friend at work.

The regression models also controlled for various demographic factors. Notably, higher education levels are negatively associated with engagement for Millennial women (β=-0.202, $p<0.01$). Longer tenure in one's career is also negatively related to engagement for Baby Boomer women (β=-0.191, $p<0.05$) and Gen X women (β=-0.140, $p<0.05$), while longer tenure in the current organization is positively linked to engagement for Millennial women (β=0.287, $p<0.01$) and men (β=0.085, $p<0.05$).

TABLE 4: MODEL 1 - ORIGINAL EMPLOYEE ENGAGEMENT STANDARIDIZED OLS REGRESSION RESULTS, BY GENERATION & GENDER

	Baby Boomer		Gen X		Millenial		Gen Z		All	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Employee Engagement Questions										
Do you know what is expected of you at work?	0.242*	-0.052	0.007	0.339**	0.169	0.341***	0.048	0.530	0.059	0.238***
Do you have the materials and equipment to do your work right?	-0.222*	0.268*	-0.160*	-0.026	-0.038	0.193*	0.747*	1.267	0.026	0.075
At work, do you have the opportunity to do what you do best every day?	0.189	-0.099	0.269**	0.103	0.073	-0.161	-0.053	-0.569	0.163**	-0.026
In the last seven days, have you received recognition or praise for doing good work?	0.225	0.026	0.214*	-0.090	-0.003	0.081	0.059	-0.367	-0.037	0.031
Does your supervisor, or someone at work, seem to care about you as a person?	0.281*	0.086	-0.091	0.121	0.222*	-0.069	0.098	1.396	0.232***	0.041
Is there someone at work who encourages your development?	-0.232	-0.104	0.132	-0.039	-0.049	0.132	-0.069	-0.364	0.021	-0.055
At work, do your opinions seem to count?	0.161	0.121	-0.099	-0.037	0.110	0.194	0.128	0.946	-0.014	0.093
Does the mission/purpose of your company make you feel your job is important?	0.278*	0.120	0.346***	0.142	0.094	0.237*	-0.123	-0.408	0.2568***	0.241***
Are your associates (fellow employees) committed to doing quality work?	-0.055	0.155	-0.127	0.062	0.138	0.016	0.013	-0.618	-0.013	0.076
Do you have a best friend at work?	0.193*	0.053	0.047	0.271**	-0.007	0.220*	0.183	-0.559	0.077	0.159**
In the last six months, has someone at work talked to you about your progress?	-0.095	0.077	-0.181*	0.038	-0.103	-0.093	0.384	0.350	-0.043	0.011
In the last year, have you had opportunities to learn and grow?	0.027	0.346*	0.413***	0.102	0.235*	0.125	-0.112	-0.567	0.136*	0.129*
Controls										
Race	0.026	0.082	-0.020	-0.135	-0.046	-0.057	-0.233	-0.032	-0.044	0.036
Ethnicity	-0.012	-0.138	-0.094	-0.017	0.030	0.061	0.233	-0.049	0.056	-0.007
State of Residence	-0.087	-0.110	0.050	0.091	-0.125	0.221*	0.063	--	-0.025	0.046
Birth Year	-0.133	0.073	-0.071	-0.064	-0.019	-0.016	0.243	-0.001	-0.261***	-0.107
Education Level	-0.005	0.023	0.015	0.002	-0.202**	-0.107	-0.001	-0.211	-0.063	-0.029
Marital Status	0.085	-0.027	0.190**	-0.061	-0.022	-0.069	0.165	0.428	0.060	-0.068
Years Worked in Career	-0.191*	0.098	-0.140*	-0.045	-0.318***	-0.079	0.020	1.047	-0.176*	-0.019
Years Worked in Current Organization	0.088	0.088	-0.008	0.104	0.287**	-0.015	0.247	-0.395	0.057	0.085*
N	61	68	95	96	109	76	35	20	300	261
Adjusted R-Squared	**0.514**	**0.519**	**0.598**	**0.452**	**0.459**	**0.566**	**0.113**	--	**0.458**	**0.4999**
F	4.17***	4.61***	7.98***	4.91***	5.58***	5.90***	1.22	--	13.64***	13.99***

*Note: Beta values; Significance levels: * p < .05; ** p < .01; *** p < .001*

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

The regression models demonstrate strong explanatory power, with adjusted R-squared values ranging from 0.452 to 0.598 for the generational cohorts. The overall models are statistically significant, with F-statistics significant at the $p<0.001$ level, indicating that the included variables can account for a substantial portion of the variance in employee engagement within each generational group. These findings underscore the importance of tailoring engagement strategies to address the unique needs and experiences of a diverse, multigenerational workforce.

Model 2: Revised Employee Engagement Model

The revised regression analysis presented in Table 5 provides a more comprehensive understanding of the factors influencing employee engagement across different generational cohorts, by gender. The regression results in this revised model continue to reveal significant generational and gender disparities in employee engagement. Among Baby Boomers, being female is positively associated with having the opportunity to do what they do best every day ($\beta=0.234$, $p<0.10$) and having a best friend at work ($\beta=0.232$, $p<0.05$). In contrast, for Generation X, being female is negatively related to having the necessary materials and equipment ($\beta=-0.091$, $p<0.10$) and receiving progress feedback ($\beta=-0.196$, $p<0.10$). For Millennials, being female is negatively linked to having opportunities to learn and grow ($\beta=0.160$, $p<0.10$). Demographic factors also play a role, as longer tenure in one's career is negatively associated with engagement for Baby Boomer women ($\beta=-0.239$, $p<0.05$), Gen X women ($\beta=-0.149$, $p<0.05$), and Millennial women ($\beta=-0.139$, $p<0.05$), while higher education levels are also negatively related to engagement for Millennial women ($\beta=-0.152$, $p<0.05$). Interestingly, being married is positively linked to engagement for Gen X women ($\beta=0.249$, $p<0.01$).

The regression analysis also examined employee activation, revealing further generational and gender differences. Baby Boomer men report a stronger sense of what makes their job meaningful ($\beta=0.432$, $p<0.05$) and higher organizational commitment ($\beta=0.350$, $p<0.05$). Gen X men exhibit a greater sense of purpose ($\beta=0.255$, $p<0.05$) and belief that their work group is where they are meant to be ($\beta=0.176$, $p<0.10$). Millennial women, on the other hand, have a stronger belief that their work group is where they are meant to be ($\beta=0.476$, $p<0.001$) and higher organizational commitment ($\beta=0.269$, $p<0.01$).

TABLE 5: MODEL 2 - REVISED EMPLOYEE ENGAGEMENT STANDARIDIZED OLS REGRESSION RESULTS, BY GENERATION & GENDER

	Baby Boomer		Gen X		Millenial		Gen Z		All	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Employee Engagement Questions										
Do you know what is expected of you at work?	0.146	0.083	0.036	0.349**	0.143	0.167	-0.209	0.147	0.069	0.206***
Do you have the materials and equipment to do your work right?	-0.216	0.081	-0.091	-0.091	0.140	0.078	0.025	-0.096	0.064	-0.017
At work, do you have the opportunity to do what you do best every day?	0.234	-0.238	0.208*	0.133	-0.060	-0.099	0.287	-0.074	0.104*	-0.022
In the last seven days, have you received recognition or praise for doing good work?	0.198	0.062	0.134	-0.127	0.000	-0.102	-0.217	0.126	-0.037	-0.037
Does your supervisor, or someone at work, seem to care about you as a person?	0.397*	0.065	-0.069	0.068	0.014	-0.032	0.318	0.963	0.173**	-0.029
Is there someone at work who encourages your development?	-0.287	-0.275	0.155	-0.017	-0.011	0.074	-0.076	--	0.007	-0.088
At work, do your opinions seem to count?	0.065	0.062	-0.036	-0.035	-0.132	0.024	0.236	0.139	-0.043	0.055
Does the mission/purpose of your company make you feel your job is important?	0.166	-0.002	0.155	0.060	-0.014	0.215	-0.301	-0.012	0.082	0.1101*
Are your associates (fellow employees) committed to doing quality work?	-0.157	-0.058	-0.155	0.042	0.063	-0.043	0.064	--	-0.024	-0.024
Do you have a best friend at work?	0.232*	-0.047	0.004	0.198*	0.035	0.064	-0.132	-0.393	0.020	0.090*
In the last six months, has someone at work talked to you about your progress?	-0.075	0.241	-0.196	0.057	-0.043	-0.012	0.228	0.424	-0.042	0.098
In the last year, have you had opportunities to learn and grow?	0.089	0.411	0.328**	0.049	0.160	0.146	-0.077	-0.400	0.056	0.098
Employee Activation Questions										
I have a good sense of what makes my job meaningful.	0.270	0.432*	0.175	-0.155	-0.207	0.219	0.453	--	0.045	0.075
I have discovered work that has a satisfying purpose.	-0.057	-0.031	0.135	0.255*	0.143	0.205	0.554	0.001	0.167**	0.192**
I believe that my work group is where I am meant to be.	-0.056	0.001	0.086	0.176	0.476***	0.205	-0.323	0.902	0.155*	0.192**
I see myself as a leader.	-0.010	-0.029	0.005	0.068	0.097	0.213	0.278	0.104	0.080*	0.055
I would be very happy to spend the rest of my career with this organization.	0.124	0.350*	-0.031	0.017	0.269**	0.029	0.092	-0.316	0.146*	0.140*
Controls										
Race	0.004	0.077	-0.066	-0.123	0.044	-0.076	0.040	--	-0.050	0.042
Ethnicity	0.021	-0.181	-0.090	-0.063	0.036	0.078	0.166	-0.588	0.047	-0.043
State of Residence	-0.021	-0.030	0.055	0.106	-0.144*	0.124	-0.249	-0.361	-0.030	0.008
Birth Year	-0.140	0.039	-0.068	-0.045	-0.026	0.076	0.106	0.052	-0.233***	-0.100
Education Level	0.118	-0.040	0.011	0.054	-0.152*	-0.096	-0.017	--	-0.059	-0.028
Marital Status	0.105	0.038	0.249**	-0.077	-0.058	0.013	0.075	0.559	0.093*	-0.007
Years Worked in Career	-0.239*	0.018	-0.149*	-0.021	-0.139*	-0.009	-0.089	0.968	-0.175**	-0.018
N	61	68	95	96	109	76	35	20	300	261
Adjusted R-Squared	**0.513**	**0.599**	**0.622**	**0.464**	**0.628**	**0.711**	**0.373**	--	**0.537**	**0.583**
F	3.61***	5.18***	7.43***	4.43***	8.60***	8.69*	1.84	--	15.47***	16.14***

*Note: Beta values; Significance levels: * p < .05; ** p < .01; *** p < .001*

The regression models demonstrate strong explanatory power, with adjusted R-squared values ranging from 0.373 to 0.711 for the generational cohorts. The overall models are statistically significant, with F-statistics

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

significant at the p<0.001 level, indicating that the included variables can account for a substantial portion of the variance in both employee engagement and activation within each generational group. These findings highlight the importance of implementing tailored strategies to address the diverse needs and experiences of a multigenerational workforce.

Model 3: Best Employee Engagement Model

Finally, as seen in Table 6 below, we took the most impactful engagement and activation variables from the last model, combined with our control variables, to create our model of best fit. Among Baby Boomers, being female is positively associated with knowing what is expected at work (β=0.267, p<0.05) and having a supervisor who seems to care about them as a person (β=0.280, p<0.05). For Generation X, being female is positively related to having the opportunity to do what they do best every day (β=0.203, p<0.05) and feeling the company's mission/purpose makes their job feel important (β=0.264, p<0.05). For Millennials, being female is positively linked to knowing what is expected at work (β=0.186, p<0.05) and having a supervisor who seems to care about them as a person (β=0.465, p<0.01). Demographic factors also play a role, as longer tenure in one's career is negatively associated with engagement for Gen X women (β=-0.142, p<0.05) and Millennial women (β=-0.182, p<0.01), while being married is positively linked to engagement for Gen X women (β=0.228, p<0.01).

The regression analysis also examined employee activation, revealing further generational and gender differences. Baby Boomer men report a stronger sense of purpose in their work (β=0.269, p<0.05). Gen X men exhibit a greater belief that their work group is where they are meant to be (β=0.144, p<0.10) and see themselves as leaders (β=0.213, p<0.01). Millennial women have a stronger belief that their work group is where they are meant to be (β=0.381, p<0.001) and higher organizational commitment (β=0.287, p<0.01). Millennial men show a greater sense of purpose (β=0.325, p<0.01) and see themselves as leaders (β=0.213, p<0.01).

As seen in Figure 3 below, the regression models demonstrate strong explanatory power, with adjusted R-squared values ranging from 0.494 to 0.719 for the generational cohorts. The overall models are statistically significant, with F-statistics significant at the p<0.001 level, indicating that the included variables can account for a substantial portion of the variance in both employee engagement and activation within each generational group. These findings highlight the importance of implementing tailored strategies to address the diverse needs and experiences of a multigenerational workforce.

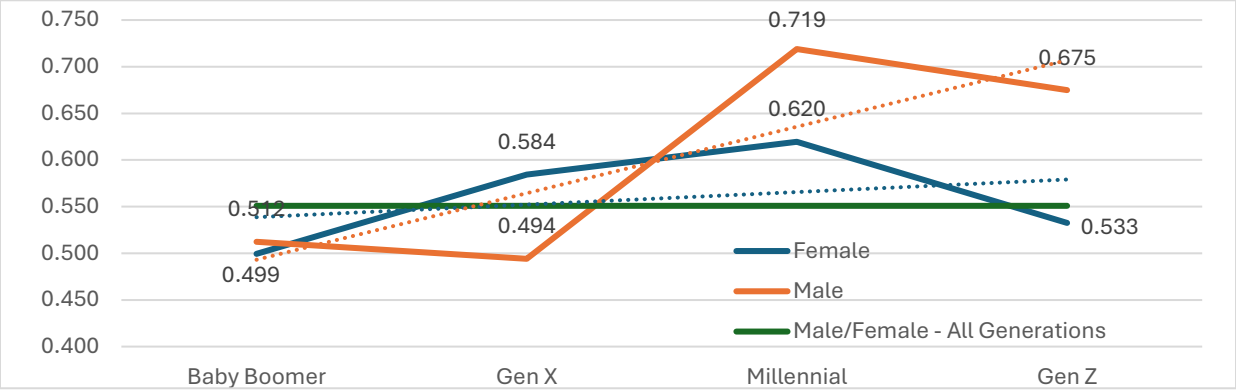
TABLE 6: MODEL 3 - BEST EMPLOYEE ENGAGEMENT STANDARIDIZED OLS REGRESSION RESULTS, BY GENERATION & GENDER

	Baby Boomer		Gen X		Millenial		Gen Z		All	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Employee Engagement Questions										
Do you know what is expected of you at work?	0.267*	0.008	-0.020	0.314**	0.186*	0.174*	-0.106	0.330	0.077	0.209***
At work, do you have the opportunity to do what you do best every day?	0.068	0.201	0.203*	0.071	-0.006	-0.066	0.224	-0.052	0.128*	0.002
Does your supervisor, or someone at work, seem to care about you as a person?	0.280*	-0.085	0.088	0.048	0.097	0.055	0.465**	0.524	0.160**	-0.012
Does the mission/purpose of your company make you feel your job is important?	0.217	0.250	0.264*	0.086	-0.025	0.211*	-0.298	0.069	0.089	0.143*
Do you have a best friend at work?	0.174	0.141	-0.016	0.233**	-0.004	0.066	-0.092	-0.185	0.015	0.113**
Employee Activation Questions										
I have discovered work that has a satisfying purpose.	0.096	0.269*	0.223*	0.171	0.047	0.325**	0.891***	0.233	0.207***	0.231***
I believe that my work group is where I am meant to be.	-0.088	-0.001	0.126	0.144	0.381***	0.308*	-0.180	0.056	0.134*	0.201**
I see myself as a leader.	0.038	0.034	0.055	0.043	0.069	0.213**	0.291*	-0.022	0.075*	0.061
I would be very happy to spend the rest of my career with this organization.	0.089	0.214	0.023	-0.041	0.287**	0.026	-0.007	0.070	0.147**	0.126*
Controls										
Race	-0.047	0.064	-0.091	-0.150	0.045	-0.047	0.276	0.354	-0.043	0.040
Ethnicity	-0.032	0.010	-0.106	-0.072	0.022	0.092	0.193	-0.209	0.047	-0.047
State of Residence	-0.071	-0.066	0.073	0.084	-0.127*	0.111	-0.388	-0.144	-0.036	0.017
Birth Year	-0.059	0.022	-0.085	-0.052	-0.030	0.089	0.233	0.121	-0.236***	-0.092
Education Level	0.057	-0.143	-0.024	0.031	-0.123*	-0.093	-0.101	0.001	-0.057	-0.022
Marital Status	0.065	-0.091	0.228**	-0.068	-0.071	0.021	0.014	0.378	0.085*	-0.022
Years Worked in Career	-0.155	-0.042	-0.142*	-0.029	-0.182**	-0.026	0.074	0.313	-0.164**	-0.035
N	61	68	95	96	109	76	35	20	300	261
Adjusted R-Squared	**0.499**	**0.512**	**0.584**	**0.494**	**0.620**	**0.719**	**0.533**	**0.675**	**0.544**	**0.582**
F	4.74***	5.40***	9.25***	6.80***	11.99***	12.99***	3.42	3.46***	23.28***	23.63***

*Note: Beta values; Significance levels: * p < .05; ** p < .01; *** p < .001*

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

FIGURE 3: MODEL ADJUSTED R-SQUARED, BY GENERATION



Variations in Model Fit—Why it Matters

The substantial variations in adjusted R-squared values across the generational and gender-specific regression models highlight the critical value of taking a nuanced, age- and gender-specific approach to understanding and promoting employee engagement. The findings clearly demonstrate that the key drivers of engagement can differ markedly not only between generational cohorts, but also within them based on gender.

By recognizing these nuanced generational and gender differences in what motivates and fulfills workers, organizations can develop far more effective, tailored strategies to foster commitment and performance across their multigenerational workforce. A one-size-fits-all approach to engagement is unlikely to be optimally effective given the diverse needs and preferences of employees at different life and career stages, as well as the unique priorities and psychological needs of men and women within each generation.

For example, the analysis revealed that discovery of meaningful, purposeful work was a particularly strong engagement driver for Baby Boomer and Generation Z men, while a sense of belonging to one's work group was most salient for Millennial women. Likewise, supervisor support emerged as exceptionally important for engaging Generation Z employees, with the effect being more pronounced for women. Clearly, the specific factors that cultivate high engagement vary considerably based on the employee's generational background and gender.

Organizations that tailor their engagement initiatives accordingly, perhaps offering a suite of options to accommodate diverse generational and gender preferences, are far more likely to effectively inspire commitment, productivity, and well-being across their multigenerational workforce. This may involve anything from restructuring recognition and rewards programs, to enhancing opportunities for skill development and growth, to fostering more personalized manager-employee relationships.

By deeply understanding the unique engagement determinants for each generational and gender cohort, companies can make strategic, data-driven investments in their human capital that maximize returns in the form of elevated performance, reduced turnover, and stronger organizational culture. This nuanced, intersectional approach represents a crucial competitive advantage in today's dynamic, multigenerational work environments.

The substantial variations in explanatory power across the age- and gender-specific regression models underscores that a one-size-fits-all mentality is ill-advised when it comes to fostering employee engagement. Organizations that recognize and respond to the nuanced generational and gender differences in work

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

motivation and fulfillment will be best positioned to cultivate a highly engaged, high-performing, and resilient multigenerational workforce.

Revisting Hypotheses

The research findings allow us to evaluate the hypotheses proposed at the outset of the study:

- *Hypothesis 1:* The results supported this hypothesis. The data revealed significant differences in employee engagement levels across the generational cohorts, with Baby Boomers reporting the highest levels of engagement, followed by Generation X, Millennials, and Generation Z. Similarly, gender differences were observed, with men generally exhibiting higher engagement levels than women, except for at the senior leadership level where the trend reversed.
- *Hypothesis 2a:* The findings supported this hypothesis. The regression analyses demonstrated that the relative importance of basic needs and individual contribution factors in predicting employee engagement differed significantly across the generational cohorts and between men and women.
- *Hypothesis 2b:* This hypothesis was partially supported. The results indicated that basic needs determinants were indeed more predictive of engagement for the younger generations (Millennials and Generation Z) compared to the older cohorts (Baby Boomers and Generation X). However, the gender differences were less pronounced, with basic needs factors being similarly important for both men and women in shaping engagement.
- *Hypothesis 2c:* This hypothesis was supported by the findings. The regression models showed that individual determinants, such as sense of meaning and purpose, were more influential in predicting engagement for the older generational cohorts (Baby Boomers and Generation X) compared to the younger generations (Millennials and Generation Z). Additionally, individual factors were more salient for men in comparison to women.
- *Hypothesis 3:* The results partially supported this hypothesis. Teamwork factors were indeed more predictive of engagement for the younger generations (Millennials and Generation Z) compared to the older cohorts. However, the gender differences were not as pronounced, with teamwork determinants being similarly important for both men and women in shaping engagement.
- *Hypothesis 4*: This hypothesis was supported by the findings. The regression analyses indicated that growth factors, such as opportunities for learning and development, were more influential in predicting engagement for the older generational cohorts (Baby Boomers and Generation X) compared to the younger generations (Millennials and Generation Z). Additionally, growth determinants were more salient for men than for women in shaping engagement.
- *Hypothesis 5:* This hypothesis was partially supported. The results showed that worker activation factors, such as purposeful work, sense of belonging, and organizational commitment, were more predictive of engagement for the younger generations (Millennials and Generation Z) compared to the older cohorts. However, the gender differences were less pronounced, with worker activation determinants being similarly important for both men and women in shaping engagement.

Overall, the findings largely supported the study's hypotheses, revealing significant generational and gender-specific differences in the drivers of employee engagement. These insights can inform the development of tailored, workforce-responsive strategies to foster commitment, performance, and well-being across an organization's diverse employee base.

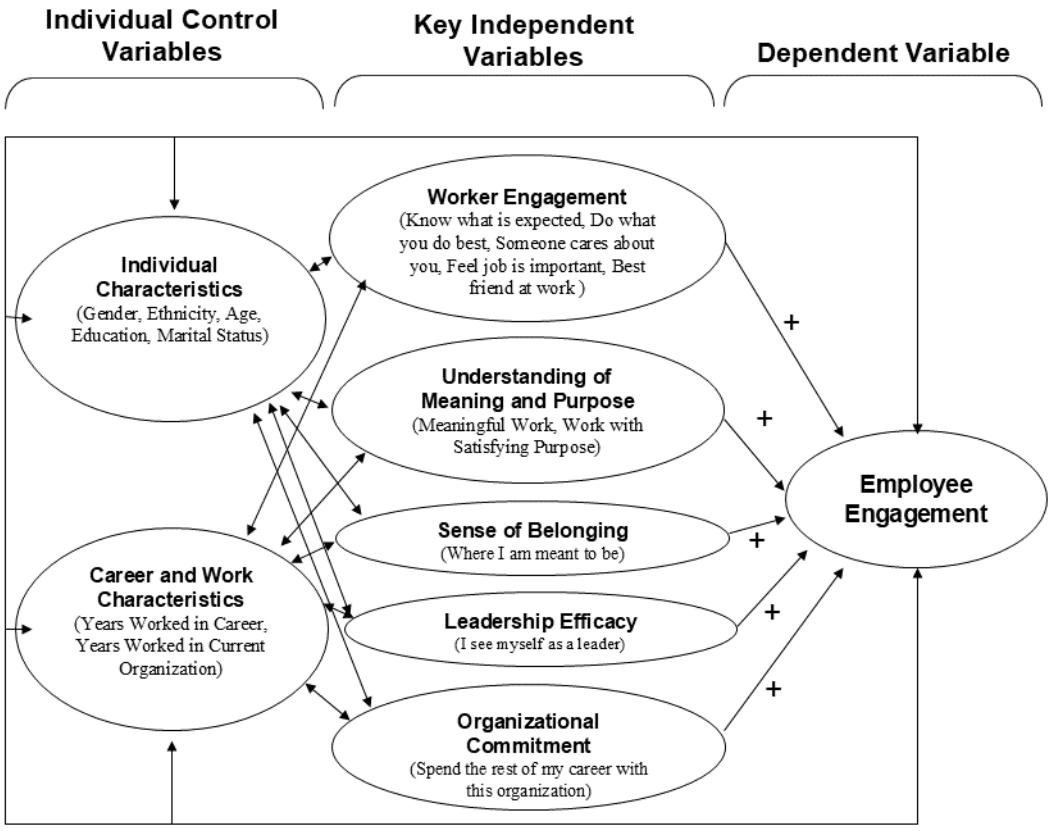
eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

A Revised Employee Engagement Model

The initial conceptual framework and associated hypotheses presented in Figure 1 only partially captured the complex relationships uncovered in this study between employee engagement, generational cohort, and key workplace factors. While traditional determinants like basic needs fulfillment, individual contributions, teamwork, and growth opportunities maintained relevance, the prominent influence of the worker activation constructs necessitated an update to the conceptualization.

The revised conceptual framework shown in Figure 4 incorporates insights from the research on generational differences in engagement drivers. Notably, it positions the worker activation dimensions - including purposeful work, sense of belonging, leadership efficacy, and organizational commitment - as core influencers of employee engagement, rather than separate supplementary predictors. By conceptualizing worker activation as a multidimensional construct composed of these factors, the updated model provides a more robust and comprehensive perspective for understanding engagement across diverse generational cohorts. This refined view recognizes the central role of worker activation in driving engagement, particularly given the evolving nature of work and shifting organizational norms.

FIGURE 4: REVISED RESEARCH MODEL



Placing worker activation at the core of the model acknowledges research demonstrating that employee engagement is influenced more by the discretionary commitment cultivated through inclusive, empowering corporate cultures, rather than solely by the fulfillment of basic expectations. The updated framework further

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

underscores the vital importance of activation in motivating discretionary effort across all age groups to achieve optimal outcomes for both individuals and organizations.

Crucially, the revised conceptual model accounts for potential variations in the salience of activation determinants based on the employee's generational background. For example, a strong sense of belonging may be a particularly potent engagement driver for Millennials, while purposeful work could be more influential for Baby Boomers. By incorporating these generational nuances, the updated framework provides a more holistic understanding of the multifaceted factors shaping engagement in today's multigenerational workforce.

This revised conceptual model presents new opportunities for maximizing diverse and thriving workforces through customized approaches tailored to nurturing high activation among employees of all ages. Rather than a fixed state, engagement may depend on specific generational contexts and be shaped not just by individual attributes, but also by strategically designed workplace experiences that adapt to evolving organizational and societal norms. This perspective can guide future theory development and ongoing study of employee engagement across different age cohorts.

DISCUSSION

The findings from this comprehensive study contribute to the growing body of research on employee engagement by providing a more nuanced, age- and gender-specific understanding of the key drivers of discretionary effort in the workplace. The results offer several important insights that can inform organizational strategies and HR practices to optimize engagement and performance across diverse generational and gender groups.

One of the key takeaways is the significant differences observed in employee engagement levels between the generational cohorts and between men and women. Consistent with previous research, Baby Boomers reported the highest levels of engagement, followed by Generation X, Millennials, and Generation Z (Costanza & Finkelstein, 2015). This pattern may be attributed to factors such as differing work values, career expectations, and life-stage considerations across the age groups (Parry & Urwin, 2011; Twenge et al., 2010). The gender differences, with men generally exhibiting higher engagement levels than women except at the senior leadership level, align with recent global surveys and may be linked to issues like isolation, lack of emotional support, and the absence of close relationships that can characterize top-level roles for women (Frumar & Truscott-Smith, 2024).

The study's regression analyses provide deeper insights into how the relative influence of various engagement determinants varies across generations and between genders. Consistent with Hypothesis 2b, basic needs factors were more salient in predicting engagement for the younger cohorts (Millennials and Generation Z) compared to older generations, suggesting that fulfilling foundational physiological, safety, and belongingness needs is particularly important for driving discretionary effort among younger workers. This finding aligns with life-stage theory, which posits that younger employees are more focused on satisfying basic needs early in their careers (Schaie, 1965; Super, 1957).

Interestingly, the gender differences in the importance of basic needs were less pronounced, with these factors being similarly predictive of engagement for both men and women. This contrasts with the hypothesis (2b) that basic needs would be more salient for women, and may indicate that the generational differences outweigh the gender-based variations in the relative influence of these foundational engagement drivers.

The results also supported the notion that individual determinants, such as sense of meaning and purpose, are more influential in shaping engagement for older generations (Baby Boomers and Generation X)

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

compared to younger cohorts (Hypothesis 2c). This finding aligns with the life-stage theory, which suggests that as employees progress in their careers, they tend to place greater emphasis on intrinsic factors like the meaningfulness of their work (Schaie, 1965; Super, 1957). The gender differences were also in line with the hypothesis, with individual determinants being more salient for men than for women in predicting engagement.

Regarding teamwork factors, the results confirmed that these determinants are more predictive of engagement for the younger generations (Millennials and Generation Z) compared to older cohorts, as hypothesized (Hypothesis 3). This may be attributed to the greater emphasis that younger employees place on collaboration, social connections, and a positive work environment (Costanza & Finkelstein, 2015; Parry & Urwin, 2011). However, the gender differences in the influence of teamwork factors were less pronounced, suggesting that the generational variations may be more significant than the gender-based differences in this regard.

The findings also supported the hypothesis (4) that growth determinants, such as opportunities for learning and development, are more salient in predicting engagement for the older generations (Baby Boomers and Generation X) compared to younger cohorts. This aligns with the notion that as employees mature in their careers, they tend to prioritize opportunities for personal and professional growth (Schaie, 1965; Super, 1957). The gender differences were also in line with the hypothesis, with growth factors being more influential for men than for women in shaping engagement.

Regarding worker activation determinants, the results partially supported the hypothesis (5), indicating that these factors (e.g., purposeful work, sense of belonging, organizational commitment) are more predictive of engagement for the younger generations (Millennials and Generation Z) compared to older cohorts. This suggests that empowering, activating work environments may be particularly important for fostering discretionary effort among younger employees. However, the gender differences were less pronounced, with worker activation determinants being similarly important for both men and women in shaping engagement.

PRACTICAL IMPLICATIONS & RECOMMENDATIONS FOR ORGANIZATIONS & WORKERS

The implications of these findings are twofold. First, they underscore the need for organizations to adopt a more tailored, generationally- and gender-responsive approach to employee engagement initiatives. Strategies that effectively engage Baby Boomers and Generation X may not resonate as strongly with Millennials and Generation Z, and vice versa. Similarly, engagement drivers that motivate men may differ from those that are most salient for women, particularly at the senior leadership level.

Second, the insights gained from this study can inform the development of comprehensive, workforce-responsive engagement strategies that address the unique needs and preferences of an organization's multigenerational, gender-diverse employee base. By aligning engagement initiatives with the distinct drivers of discretionary effort across age and gender groups, organizations can foster stronger commitment, performance, and well-being throughout their diverse workforce.

Some potential strategies include:

1. Tailored onboarding and development programs that cater to the basic needs, individual growth aspirations, teamwork preferences, and activation requirements of different generational cohorts.
2. Flexible work arrangements, leadership opportunities, mentorship programs, and inclusive language training to support women's engagement and advancement, particularly in senior roles.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

3. Emphasis on meaningful work, autonomy, and opportunities for continuous learning and development to engage older employees.
4. Collaborative work environments, strong team relationships, and clear communication of organizational purpose and values to foster engagement among younger generations.
5. Comprehensive employee listening and feedback mechanisms to capture the evolving needs and preferences of the multigenerational, gender-diverse workforce.

By adopting these generationally- and gender-responsive approaches, organizations can optimize employee engagement and discretionary effort, ultimately driving improved organizational performance and outcomes.

Overall, this study's findings underscore the importance of understanding the nuanced, age- and gender-specific determinants of employee engagement. By tailoring engagement strategies to the unique needs and preferences of an organization's diverse workforce, leaders and HR professionals can cultivate a highly committed, productive, and resilient employee base - a critical competitive advantage in today's rapidly evolving business landscape.

LIMITATIONS & OPPORTUNITIES FOR FUTURE RESEARCH

While this study offers valuable insights into the generational and gender-based differences in employee engagement drivers, it is important to acknowledge several limitations that provide opportunities for future research.

First, the cross-sectional nature of the data limits the ability to make causal inferences about the relationships between the various engagement determinants and employee outcomes. Longitudinal research designs would allow for a deeper understanding of how these dynamics evolve over time, particularly as generational cohorts and gender representation in the workforce continue to shift.

Second, the data was collected from a sample of U.S. employees, which may limit the generalizability of the findings to other cultural and geographical contexts. Replicating this study in different countries and regions would provide a more comprehensive, globally-representative understanding of generational and gender-based engagement patterns.

Third, the study focused on the main effects of generational cohorts and gender on engagement drivers, but did not explore potential interaction effects. It is possible that the influence of certain determinants may be amplified or diminished when considering the intersections between age and gender. Future research should investigate these more nuanced, interactive effects to uncover a richer, more holistic picture of engagement dynamics.

Fourth, the analysis was limited to the broad generational categories commonly used in organizational research (Baby Boomers, Generation X, Millennials, Generation Z). While this approach aligns with established frameworks, it may overlook important within-group variations that could provide additional insights. Examining engagement drivers at a more granular, year-of-birth level may reveal more fine-grained differences across the generational spectrum.

Fifth, the study relied on self-reported survey data, which may be subject to common method bias and social desirability effects. Incorporating multi-source data, such as manager assessments, performance metrics, and objective organizational records, could enhance the validity and reliability of the findings.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Finally, the current study focused primarily on the main drivers of engagement, but did not explore potential mediating or moderating mechanisms that could help explain the observed generational and gender differences. Future research should investigate the underlying psychological, social, and organizational processes that shape engagement outcomes for different age and gender groups.

Despite these limitations, this study offers a strong foundation for continued exploration of generational and gender-based engagement dynamics. Opportunities for future research include:

1. Longitudinal investigations to examine the evolving nature of engagement drivers over time, particularly as workforce demographics continue to shift.
2. Cross-cultural studies to understand how cultural and national contexts may influence the generational and gender-based patterns observed in this research.
3. Intersectional analyses to uncover the interactive effects of age, gender, and other demographic factors on engagement determinants.
4. Granular, year-of-birth level examinations of generational differences to identify more nuanced variations within broad age cohorts.
5. Multi-source data collection and mixed-methods approaches to enhance the validity and depth of the findings.
6. Exploration of mediating and moderating mechanisms that can help explain the underlying processes shaping engagement outcomes for diverse generational and gender groups.

By addressing these limitations and pursuing these research directions, scholars can build upon the insights provided by this study to further advance the understanding of employee engagement in the context of an increasingly diverse, multigenerational workforce. These efforts can inform the development of tailored, evidence-based strategies to foster commitment, performance, and well-being across organizations.

CONCLUSION

This comprehensive study provides valuable insights into the generational and gender-based differences in the drivers of employee engagement. By analyzing survey data from over 500 U.S. employees, the research offers a nuanced understanding of how the relative influence of various engagement determinants varies across age cohorts and between men and women.

The findings reveal significant differences in engagement levels, with Baby Boomers reporting the highest levels, followed by Generation X, Millennials, and Generation Z. Gender differences were also observed, with men generally exhibiting higher engagement than women, except at the senior leadership level where the trend reversed.

The regression analyses shed light on the distinct engagement drivers for different generational and gender groups. Basic needs factors, such as physiological and safety requirements, were more salient in predicting engagement for the younger cohorts (Millennials and Generation Z) compared to older generations. In contrast, individual determinants, like sense of meaning and purpose, were more influential for the older cohorts (Baby Boomers and Generation X) and for men compared to women.

Teamwork-related factors were more predictive of engagement for the younger generations (Millennials and Generation Z), while growth determinants, including opportunities for learning and development, were more important for the older cohorts (Baby Boomers and Generation X) and for men. Worker activation variables, such as purposeful work and organizational commitment, were more salient in

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

shaping engagement for the younger generations compared to older cohorts, but exhibited similar importance for both men and women.

These findings underscore the need for organizations to adopt a more tailored, generationally- and gender-responsive approach to employee engagement initiatives. By aligning engagement strategies with the distinct drivers of discretionary effort across age and gender groups, leaders and HR professionals can foster stronger commitment, performance, and well-being throughout their diverse workforce.

Potential strategies include tailored onboarding and development programs, flexible work arrangements and inclusive leadership opportunities for women, emphasis on meaningful work and continuous learning for older employees, and collaborative work environments and clear communication of organizational purpose for younger generations.

Future research directions include longitudinal investigations, cross-cultural studies, intersectional analyses, and the exploration of underlying mechanisms that shape engagement outcomes for diverse generational and gender groups. By building on the insights provided by this study, scholars can continue to advance the understanding of employee engagement in the context of an increasingly diverse, multigenerational workforce.

Overall, this research offers critical guidance for organizations seeking to optimize engagement and discretionary effort within their multigenerational, gender-diverse employee base - a strategic imperative for driving improved organizational performance and sustained competitive advantage in the modern business landscape.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

References

Albrecht, S.L. (2010) Handbook of employee engagement: Perspectives, issues, research and practice. Edward Elgar Publishing.

Bakker, A.B. and Demerouti, E. (2008) 'Towards a model of work engagement', Career Development International, 13(3), pp. 209-223.

Bakker, A.B., Demerouti, E. and Sanz-Vergel, A.I. (2014) 'Burnout and work engagement: The JD-R approach', Annual Review of Organizational Psychology and Organizational Behavior, 1(1), pp. 389-411.

Becton, J.B., Walker, H.J. and Jones-Farmer, A. (2014) 'Generational differences in workplace behavior', Journal of Applied Social Psychology, 44(3), pp. 175-189.

Breevaart, K., Bakker, A., Hetland, J., Demerouti, E., Olsen, O.K. and Espevik, R. (2014) 'Daily transactional and transformational leadership and daily employee engagement', Journal of Occupational and Organizational Psychology, 87(1), pp. 138-157.

Christian, M.S., Garza, A.S. and Slaughter, J.E. (2011) 'Work engagement: A quantitative review and test of its relations with task and contextual performance', Personnel Psychology, 64(1), pp. 89-136.

Costanza, D.P. and Finkelstein, L.M. (2015) 'Generationally based differences in the workplace: Is there a there there?', Industrial and Organizational Psychology, 8(3), pp. 308-323.

Demerouti, E., Bakker, A.B., Nachreiner, F. and Schaufeli, W.B. (2001) 'The job demands-resources model of burnout', Journal of Applied Psychology, 86(3), pp. 499-512.

Frumar, C. and Truscott-Smith, A. (2024) 'Women's engagement advantage disappears in leadership roles', Gallup Workplace, 9 July.

Gallup (2022) 2022 State of the Global Workplace.

Hackman, J.R. and Oldham, G.R. (1976) 'Motivation through the design of work: Test of a theory', Organizational Behavior and Human Performance, 16(2), pp. 250-279.

Halbesleben, J.R. (2010) 'A meta-analysis of work engagement: Relationships with burnout, demands, resources, and consequences', in Bakker, A.B. and Leiter, M.P. (eds.) Work engagement: A handbook of essential theory and research. Psychology Press, pp. 102-117.

Harter, J.K., Schmidt, F.L. and Hayes, T.L. (2002) 'Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis', Journal of Applied Psychology, 87(2), pp. 268-279.

Hartman, R.L. and Barber, E.G. (2020) 'Women in the workforce: The effect of gender on occupational self-efficacy, work engagement and career aspirations', Gender in Management, 35(1), pp. 92-118.

Hobfoll, S.E. (1989) 'Conservation of resources: A new attempt at conceptualizing stress', American Psychologist, 44(3), pp. 513-524.

Kahn, W.A. (1990) 'Psychological conditions of personal engagement and disengagement at work', Academy of Management Journal, 33(4), pp. 692-724.

Macey, W.H. and Schneider, B. (2008) 'The meaning of employee engagement', Industrial and Organizational Psychology, 1(1), pp. 3-30.

eISSN: 3143-8148 (online) doi.org/10.70175/futureofworkjournal.2026

Mannheim, K. (1952) 'The problem of generations', in Kecskemeti, P. (ed.) Essays on the sociology of knowledge. Routledge, pp. 276-320.

Maslow, A.H. (1943) 'A theory of human motivation', Psychological Review, 50(4), p. 370.

May, D.R., Gilson, R.L. and Harter, L.M. (2004) 'The psychological conditions of meaningfulness, safety and availability and the engagement of the human spirit at work', Journal of Occupational and Organizational Psychology, 77(1), pp. 11-37.

Nobles, C. (2023) 'The gender inequality statistic you might not be tracking', Achievers, 7 February.

Parry, E. and Urwin, P. (2011) 'Generational differences in work values: A review of theory and evidence', International Journal of Management Reviews, 13(1), pp. 79-96.

Pew Research Center (2019) Defining generations: Where Millennials end and Generation Z begins.

Rich, B.L., Lepine, J.A. and Crawford, E.R. (2010) 'Job engagement: Antecedents and effects on job performance', Academy of Management Journal, 53(3), pp. 617-635.

Ryan, R.M. and Deci, E.L. (2000) 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being', American Psychologist, 55(1), pp. 68-78.

Saks, A.M. (2006) 'Antecedents and consequences of employee engagement', Journal of Managerial Psychology, 21(7), pp. 600-619.

Sharma, A., Goel, A. and Sengupta, S. (2017) 'How does work engagement vary with employee demography?: Revelations from the Indian IT industry', Procedia Computer Science, 122, pp. 146-153.

Schaie, K.W. (1965) 'A general model for the study of developmental problems', Psychological Bulletin, 64(2), pp. 92-107.

Schaufeli, W.B. and Bakker, A.B. (2004) 'Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study', Journal of Organizational Behavior, 25(3), pp. 293-315.

Schaufeli, W.B., Salanova, M., González-Romá, V. and Bakker, A.B. (2002) 'The measurement of engagement and burnout: A two-sample confirmatory factor analytic approach', Journal of Happiness Studies, 3(1), pp. 71-92.

Shuck, B. and Wollard, K. (2010) 'Employee engagement and HRD: A seminal review of the foundations', Human Resource Development Review, 9(1), pp. 89-110.

Strauss, W. and Howe, N. (1991) Generations: The history of America's future, 1584 to 2069. William Morrow & Company.

Super, D.E. (1957) The psychology of careers. Harper & Row.

Twenge, J.M., Campbell, S.M., Hoffman, B.J. and Lance, C.E. (2010) 'Generational differences in work values: Leisure and extrinsic values increasing, social and intrinsic values decreasing', Journal of Management, 36(5), pp. 1117-1142.

Xanthopoulou, D., Bakker, A.B., Demerouti, E. and Schaufeli, W.B. (2007) 'The role of personal resources in the job demands-resources model', International Journal of Stress Management, 14(2), p. 121.

Westover, J.H. and Andrade, M.S. (2024) 'Examining the influence of employee activation on gender differences in employee engagement', Journal of Business Diversity, 24(3), pp. 78-95.

Zoe Talent Solutions (2024) Breakdown of male vs. female employee engagement statistics.

