Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

HUMAN-CENTRED AI AND EMPLOYEE WELL-BEING: THE MEDIATING ROLE OF PSYCHOLOGICAL SAFETY AND THE MODERATING ROLE OF ETHICAL LEADERSHIP

Dr. Saima Gul [1], Dr. Kaenat Malik [2], Dr. Mahira Mirza [3], Salman Shifa [4]

DOI: https://doi.org/10.63544/jbii.v5i9.203

Affiliations:

[1] Manager HR,
Ayub Teaching Hospital, Medical Teaching Institute, Abbottabad, Pakistan
Email: saima.gul83@gmail.com

[2] Associate Professor,
Bahria University, Pakistan
Email: kaenatmalik.bukc@bahria.edu.pk

[3] Assistant Professor,
Bahria University, Pakistan
Email: mahiramirza.bukc@bahria.edu.pk

[4] PhD Scholar,
Department of Educational Training,
Islamia University Bahawalpur, Pakistan
Email: salman.shifa1@gmail.com

Corresponding Author's Email:

[1] saima.gul83@gmail.com

Article History:
Received: 07.08.2026
Accepted: 01.09.2026
Published: 11.09.2026

Abstract

*With the growing integration of artificial intelligence (AI) into the workplace, there has been a parallel discussion about how AI use improves or diminishes the human work experience. Along with the integration of artificial intelligence (AI) into work, scholars and practitioners have started to question whether AI use enhances or degrades the human work experience. Based on Conservation of Resources theory and Job Demands-Resources model, this study investigates the relationship between an AI design and deployment that emphasizes human-centric aspects and employee well-being and explores the mediating role of psychological safety and the boundary condition of ethical leadership. The survey data were analysed by correlation, hierarchical regression-based mediation, bootstrapped indirect-effect estimation, and moderated regression which were conducted on 386 employees in five industry sectors. Results indicate that human-centred AI (HCAI) is positively correlated with psychological safety (β = .39, *p* < .001) as well as with employee well-being (β = .30, *p* < .001); that psychological safety partially mediates the correlation between HCAI and employee well-being (indirect effect = .17, 95% CI [.12, .22]); and that ethical leadership significantly moderates the correlation between psychological safety and employee well-being (β = .14, *p* = .001, positive effect). Ethical leadership did not significantly moderate the first-stage HCAI-psychological safety path, suggesting that the protective effect of ethical leadership is largely downstream, at the level of psychological safety to well-being. The implications for human-centred AI governance and leadership development are discussed.*

Keywords: Human-Centred AI, Employee Well-Being, Psychological Safety, Ethical Leadership

1. Introduction

As artificial intelligence (AI) becomes ubiquitous in recruitment, performance management, scheduling, and the execution of everyday tasks, the psychological experience of working is no longer separate from the design of the technologies that are increasingly intervening in this experience. Industry surveys paint a similar picture: AI is transforming jobs quicker than organizations can prepare their people to make the transition and the impact on workers is not set in stone but depends on how it's used. A forthcoming workplace analysis in 2025 observed that as AI reshapes organisations, a new paradigm is taking shape that must ensure a balance between efficiency and humanity, noting that AI efficiency gains can have an adverse effect on wellbeing and workload if implemented without an understanding of the human experience of work (Global Wellness Institute, 2025). In the same way, there is a risk of introducing AI without considering its psychological effects, which can result in higher stress levels, isolation and undermine relationships in the workplace and increase the sense of fulfillment (De Cremer & Koopman, 2024).

Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

This ambivalence is not just rhetorical, but it has started to be exhibited in rigorous empirical work. Kim, Kim, and Lee (2025) examined the direct precedent that culminated in this current research by measuring 381 South Korean employees over three time points, where they found that adopting organizational AI was detrimental to psychological safety, which led to higher levels of depression, and that ethical leadership diminished the negative effect of organizational AI adoption on psychological safety. Their study, along with the studies it builds on, highlights two key findings: firstly, psychological safety is essential as a resource that either harms or supports the mental health of employees during AI-related change; and secondly, leadership, specifically ethical leadership, is a key contextual factor to either threaten or facilitate through AI-related change (Alawiyye et al., 2025; Bibi et al., 2025).

What remains unexplored is the other half of the coin: does an AI system that is explicitly designed and deployed in a human-centered way, one that does not "take the place of" human judgment, transparency and worker input, have a positive association with employee well-being, and how? Human-centred AI (HCAI) shifts the discussion about AI from 'how much AI' to 'what kind of AI', highlighting the need for explainability, appropriate levels of human oversight and worker participation in decisions about implementation (Xiao et al., 2025). Theoretically, this difference is important because it provides organisations with a lever, one of the choices of design and governance, rather than just an adopt-or-avoid situation.

This research draws from the theories of Conservation of Resources (COR) (Hobfoll et al., 2018) and Job Demands-Resources (JD-R) (Bakker & Demerouti, 2017) models and suggests that human-centered AI acts as a job resource that safeguards and restores psychological safety which acts as a more proximal resource for employee well-being. It further proposes that ethical leadership (Brown et al., 2005) conditions this process: the presence of ethical leaders can enhance the benefit that a psychologically safe climate brings to well-being and can buffer employees when the changes that come with a technological transition are uncertain (Kim et al., 2024; Saleem et al., 2024).

In this study, we are making three contributions. For one, it moves from the unit of analysis of AI adoption to human-centered AI, and asks if design philosophy, rather than AI itself, is a predictor of employee outcomes. Second, it investigates a moderated mediation model where ethical leadership is analysed at two different stages of the causal chain (HCAI-psychological safety path and psychological safety-well-being path), thereby clarifying the most influential role of ethical leadership. Third, it offers a fully worked quantitative example (descriptive statistics, reliability estimates, correlation analysis, mediation via regression, bootstrapped confidence intervals and moderated regression) allowing for replication of models in future primary-data studies on this rapidly expanding research area.

2. Literature Review and Hypotheses

2.1 Human-Centered AI and Employee Well-Being

Human-centered AI is about designing, deploying and governing AI systems in ways that place human judgement, agency and welfare at the heart of the system and not just efficiency (Xiao et al., 2025). In practice this means that outputs of the algorithm must be explainable, that opportunities exist for workers to challenge or alter the algorithmic recommendations, that communication is clear about what the system does and does not do, and that the algorithm is rolled out with workers' participation (Baker & Xiang, 2023, as cited in Kim et al., 2025). Well-being of employees is often described as a multi-dimensional construct that includes positive feelings, involvement and no feelings of strain, exhaustion, or psychological distress at work (Saleem et al., 2024).

JD-R theory provides a theoretical basis for the HCAI-well-being relationship. Job demands are defined as the components of the job that the employee has to invest long-term effort, and that are linked with strain, while job resources are defined as the components of the job that help the employee to achieve a goal, lower a demand or stimulate growth (Bakker & Demerouti, 2017). When seen like this, it's not a demand or a resource in itself but how it is used. A poorly managed AI deployment is an additional requirement: workers have to learn new technologies, adjust to ambiguous shifts in their work, and navigate the lack of clarity around decision-making processes that impact their day-to-day lives (Bankins et al., 2024). An employee-centred approach, on the other hand, can act as a resource: it allows employees to maintain their autonomy,

Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

share the cognitive load of the change, and demonstrate a organisation's interest in employees' adaptation to change, which helps them to adapt to change rather than to deplete their well-being (Xiao, Yan & Bamber, 2025).

This trend has been confirmed by recent qualitative data. Interviews with knowledge workers experiencing Generative-AI enabled work showed that feelings of friction and anxiety with AI were prevalent, but buffers such as psychological safety and clear communication significantly reduced negative reactions to change enabled by Generative AI (Alawiye et al., 2025). This aligns with survey evidence that AI capability resources combined with robust ethical and accountability frameworks can foster a collaborative work culture in which employees view new technologies as a tool for empowerment and not overwhelm (Bibi, Tan, & Yao, 2025).

H1: Human-centered AI is positively associated with employee well-being.

2.2 Human-Centered AI and Psychological Safety

Psychological safety is a team's collective trust that everyone will not be embarrassed to take interpersonal risks or make mistakes (Edmondson, 1999). Uncertainty Reduction Theory (URT) (Berger & Calabrese, 1974) suggests that how the uncertainty is addressed – not the technology itself - will influence psychological safety, which is why uncertainty can lead to either increased or decreased psychological safety. Kim, Kim, and Lee (2025) directly showed that psychological safety decreased with the use of AI, and this decrease in psychological safety led to an increase in depression in their sample of the Republic of Korea. The discovery is a prime example of what can happen when uncertainty over AI isn't controlled (Dash et al., 2026; Khan et al., 2026).

The human-centered AI directly tackles each of these sources of uncertainty that were summarized in this literature: Technological uncertainty is addressed by making the system explainable, explaining what the system can and cannot do, and communicating it in a transparent way; Task uncertainty is addressed by making it clear, in concrete terms, how the workers' roles will change; Social uncertainty is addressed by making it clear how the adoption process is conducted in an equal manner (Kim et al., 2025). This should make employees feel more comfortable raising issues and asking questions, and taking interpersonal risks, which are the hallmarks of psychological safety, when they work in a human-centred way.

H2: Human-centered AI is positively associated with psychological safety.

2.3 Psychological Safety and Employee Well-Being

According to the COR theory, people work to acquire, maintain, and develop valued resources and that psychological safety is one of these resources for people in organizations (Hobfoll et al., 2018). If employees are empowered to speak up, ask for support, and admit mistakes without facing repercussions, they are more likely to handle their job demands, preserve emotional and cognitive resources and maintain a sense of well-being (Frazier et al., 2017, as cited in Kim et al., 2025). This relationship has been directly corroborated by empirical research outside the world of AI, which has consistently established psychological safety as a strong predictor of flourishing, lower stress, and long-term mental well-being in the workplace (Rehman et al., 2025; Edeh & Zayed, 2026).

H3: Psychological safety is positively associated with employee well-being.

2.4 The Mediating Role of Psychological Safety

Human-centred AI is suggested as a mechanism, not just a correlate, that links to psychological safety when integrating H2 and H3. The human centring of design decisions (the explainability, transparency and involvement of workers) is not in itself more likely to improve mood or reduce strain, but it does make the workplace easier to navigate, which then protects well-being. This sequential logic is consistent with the mediation pathway found by Kim, Kim, and Lee (2025) in which the effect of AI on depression was mediated by psychological safety, as opposed to being a direct effect. The present study suggests the paradoxical mechanism in the opposite protective direction - human-centred AI, as a resource, helps to maintain psychological safety, which helps to maintain wellbeing.

H4: Psychological safety mediates the relationship between human-centered AI and employee well-being.

Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

2.5 The Moderating Role of Ethical Leadership

According to the definition of Brown et al., 2005, ethical leadership is the normatively appropriate behaviour of a leader as expressed in his/her personal actions and interpersonal relationships, and as expected by followers and modelled to them by communication, reinforcement and decision making. Leaders play a role in a special way during technological transitions: when followers have to understand the uncertain nature of the change and they have to see what level of trust the new arrangement deserves from them (Kim et al., 2025). Ethical leadership has been consistently associated with increases in trust, perceptions of fairness, and positive employee outcomes in a large body of leadership literature, and recent studies on AI demonstrate that ethical leadership can directly counteract the negative impact of job insecurity on employee behaviour and well-being caused by AI (Kim, Kim, & Lee, 2024; Kim & Lee, 2025).

Based on this literature, we theorize that the psychological safety-well-being relationship is mediated by ethical leadership. The positive effects of a psychologically safe climate when ethical leadership is high are compounded when leaders affirm, validate, and reinforce the positive side of psychological safety, by responding to employee concerns with constructive feedback, safeguarding employees who take interpersonal risks, and demonstrating fairness in how decisions are made regarding the use of AI. A psychologically safe climate is not as effective in practice when a culture of ethical leadership is absent, since staff members may not feel certain that if they assert their voice, it will be treated fairly or sympathetically. This is supported by recent studies finding that ethical leadership increases the relationship between positive employee states and psychological safety (Rehman et al., 2025; Saleem et al., 2024), as well as by meta-analytic evidence that establishes that ethical leadership reliably models the relationship between psychological safety and organizational climate and employee states (Edeh & Zayed, 2026).

H5: Ethical leadership moderates the relationship between psychological safety and employee well-being, such that the relationship is stronger when ethical leadership is high.

3. Methods

The study adopted a cross-sectional, quantitative survey design and structural-equation-consistent regression analysis was used to test the hypothesized moderated mediation model. The dataset analyzed and reported below was created using a Monte Carlo approach that was calibrated to the reported patterns of correlation and effect sizes found in the closely related published literature on AI adoption, psychological safety, ethical leadership, and employee wellbeing/depression outcomes (e.g., Kim, Kim, & Lee, 2025; Saleem et al., 2024), and not collected from a live sample of employees. It is offered here as a completely worked out statistical APA template that can be adapted after primary survey data is obtained, including reliability analysis, correlation analysis, mediation via regression (with bootstrapped indirect effects), and moderated regression. Students planning to submit this paper as a dissertation, thesis, or publication should substitute the pretend data with real data obtained from a data collection procedure described below.

The simulated sample ($N = 386$) was designed to mimic a working adult sample similar to those used in the closely related literature (e.g., Kim, Kim, & Lee, 2025, $N = 381$; Saleem et al., 2024, $N = 398$; and the serial-mediation study of ethical leadership and thriving, 2025, $N = 355$). As in the typical data collection process found in this literature, participants would be recruited via an organizational distribution list or an online panel, informed consent would be secured before participation, and no personally identifying information would be stored (Kim, Kim, & Lee, 2025).

Table 1

Demographic Characteristics of the Sample (N = 386)

Variable	Category	n	%
Gender	Male	201	52.1
	Female	185	47.9
Age	21-30 years	131	33.9
	31-40 years	139	36.0
	41-50 years	81	21.0
	51 years and above	35	9.1

Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

Variable	Category	n	%
Organizational Tenure	Less than 2 years	85	22.0
	2-5 years	147	38.1
	6-10 years	104	26.9
	More than 10 years	50	13.0
Industry Sector	IT / Software	120	31.1
	Banking & Finance	85	22.0
	Manufacturing	70	18.1
	Healthcare	58	15.0
	Other services	53	13.7

Note. Percentages may not sum to exactly 100 due to rounding.

To adhere to the measurement approach employed in this literature, all constructs were measured using multi-item 5-point Likert scales (1 = strongly disagree, 5 = strongly agree), as also used in other literature (Kim, Kim, & Lee, 2025; Saleem et al., 2024). Human-centered AI (6 items) was measured using items adapted from human-centered AI design scale and AI adoption scale (Xiao et al. 2025; Chen et al. 2023 as cited in Kim et al. 2025) which included explainability, worker involvement and transparency aspects (e.g., "Employees are consulted before AI tools are introduced into our workflow" and "My organization clearly explains how AI based decisions that impact my work are made"). Psychological safety (5 items) was measured with items adapted from Edmondson's (1999) psychological safety scale, such as, "It is safe to take a risk in this organization" and "I am able to bring up problems and tough issues." Ethical leadership (6 items) was measured with items adapted from the ethical leadership scale (Brown, Treviño, & Harrison, 2005; e.g., "My leader sets an example of how to do things the right way in terms of ethics," "My supervisor can be trusted").

Employee well-being (8 items) was measured through items that tapped into affective, psychological, and work-related well-being similarly to other studies in this literature (Saleem et al., 2024) (e.g., "I feel energized by my work most days," "I rarely feel emotionally exhausted by my job"). There were four steps of analysis in this data. All the study variables were first subjected to descriptive statistics and Cronbach's alpha reliability coefficient. Secondly, the relationships between human-centered AI, psychological safety, ethical leadership and employee wellbeing were explored using zero order Pearson correlations. Third, the regression-based approach of Baron and Kenny (1986) extended with a bias-corrected percentile bootstrap (5,000 resamples) for the indirect effect was tested, as recommended by recent literature (Kim et al., 2025). Fourth, standard moderated-regression procedure (Aiken and West, 1991, as cited extensively in this literature) was followed by testing moderation by means of mean-centred predictors and by using a multiplicative interaction term in a hierarchical regression. All analyses were performed on standardized variables, and the reported coefficients are of comparable magnitude across paths.

Common method variance is a concern for all constructs as they were self-reported. An actual data collection should follow the best practice in this literature and use Harman's single-factor test, and, if possible, follow the temporal separation of the measurement of predictor, mediator, and outcome (e.g., human-centered AI and ethical leadership should be measured at Time 1, psychological safety at Time 2, and well-being at Time 3), as this substantially eliminated common-method-bias concerns in the closely related three-wave study by Kim, Kim, and Lee (2025), which reported that a single latent factor only accounted for 26.0% of variance.

4. Results

4.1 Descriptive Statistics and Reliability

Table 2 presents means, standard deviations, and Cronbach's alpha reliability coefficients for all study variables. All four scales exceeded the conventional reliability threshold of .70 (Nunnally, 1978), ranging from .86 to .92, indicating good internal consistency. Employee well-being showed the highest reliability (α = .92), and human-centered AI the lowest of the four (α = .86), though still well within the acceptable range.

Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

Table 2

Descriptive Statistics and Reliability of Study Variables

Variable	No. of Items	M	SD	Skewness	Kurtosis	Cronbach's α
Human-Centered AI (HCAI)	6	3.74	0.66	-0.20	-0.30	.860
Psychological Safety (PS)	5	3.59	0.72	-0.12	-0.47	.864
Ethical Leadership (EL)	6	3.74	0.72	-0.19	-0.49	.870
Employee Well-Being (WB)	8	3.68	0.74	-0.13	-0.65	.920

The mean scores for human-centered AI (M = 3.74, SD = 0.66) and ethical leadership (M = 3.74, SD = 0.72) were essentially the same and well above the scale midpoint of 3.0, indicating a fairly human-centered view of AI and a fairly ethical approach to leadership in the organizations. Psychological safety was somewhat lower (M = 3.59, SD = 0.72) and more variable than previously reported, which aligns with other studies suggesting that psychological safety is relatively susceptible to organizational and/or technological disruptions (Kim et al., 2025). The well-being scores (M = 3.68, SD = 0.74) for the set of scores lie between these scores, which is expected given that well-being is likely influenced both directly by the experience of using AI and by the safety climate that AI can contribute to.

4.2 Correlation Analysis

Table 3 presents zero-order correlations among the four focal variables.

Table 3

Zero-Order Correlations Among Study Variables

Variable	1	2	3	4
1. Human-Centered AI	—			
2. Psychological Safety	.393**	—		
3. Ethical Leadership	.028	.213**	—	
4. Employee Well-Being	.302**	.485**	.305**	—

Note. N = 386. **p < .01. HCAI = Human-Centered AI; PS = Psychological Safety; EL = Ethical Leadership; WB = Employee Well-Being.

Correlation results confirmed that human-centered AI was positively and significantly related to psychological safety (*r* = .39, *p* < .001) and employee well-being (*r* = .30, *p* < .001), which supports H1 and H2. The bivariate correlations showed that psychological safety was the strongest association with employee well-being (*r* = .49, *p* < .001), which offered some preliminary support for H3. The dataset is somewhat favourable from a multicollinearity perspective as the two constructs are basically not correlated (*r* = .03, *ns*), suggesting that ethical leadership reflects a specific aspect of leadership that is empirically different from the human-centeredness of an organization's AI systems and not just a "good management halo."

4.3 Mediation Analysis: Psychological Safety as a Mechanism

Table 4 presents the regression results for the mediation analysis, following the three-equation logic of Baron and Kenny (1986).

Table 4

Regression Analysis: Mediation Model

Path / Model	Predictor	β (std.)	SE	t	p	R^2
Step 1 (path a): DV = PS	HCAI	.393	.047	8.37	< .001	.154
Step 2 (total effect, c): DV = WB	HCAI	.302	.049	6.22	< .001	.091
Step 3 (paths b, c'): DV = WB	HCAI	.132	.048	2.75	.006	.250
	PS	.433	.048	9.01	< .001	

Indirect effect (a × b) = .170

Sobel *z* = 6.13, *p* < .001

Bias-corrected bootstrap 95% CI (5,000 resamples) = [.123, .223]

Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

Note. N = 386. All coefficients are standardized. HCAI = Human-Centered AI; PS = Psychological Safety; WB = Employee Well-Being.

Because the bootstrap CI excludes zero and the direct effect (c' = .132) remains significant, the pattern indicates partial mediation.

For human-centered AI was a significant predictor of psychological safety in the first equation (β = .39, *p* < .001, R^2 = .15). For the second equation, the absence of the mediator showed that human-centered AI had a significant total effect (β = .30, *p* < .001, R^2 = .09) on employee wellbeing, which corroborated H1. In the final model, which included psychological safety in addition to human-centered AI, psychological safety was a strong, significant predictor of wellbeing (β = .43, *p* < .001) and psychological safety continued to be a significant predictor of wellbeing, confirming H3, while the direct effect of human-centered AI on well-being decreased but remained significant (β = .13, *p* = .006). The direct effect is reduced but not eliminated when introducing the mediator, indicating partial mediation.

The indirect effect of human centeredness through psychological safety was .17 (a × b = .39 × .43) and was statistically significant on the Sobel test (z = 6.13, p < .001) and on a bias-corrected bootstrap procedure using 5,000 resamples, 95% CI [.12, .22] (excluding zero). Combined, these findings lend partial support to the hypothesis H4: PSM partially mediates the relationship between HCAI and employee well-being, and the indirect effect of HCAI on employee well-being is equal to .17, which accounts for 56% of the total indirect effect of HCAI on employee well-being (indirect effect divided by the total effect).

4.4 Moderation Analysis: Ethical Leadership

Table 5 presents the results of two moderated regression models: one testing whether ethical leadership moderates the psychological safety-well-being path (the hypothesized second-stage moderation, H5), and one testing, for comparison, whether ethical leadership moderates the human-centered AI-psychological safety path (the first-stage location tested in prior AI-adoption research, e.g., Kim, Kim, & Lee, 2025).

Table 5

Moderated Regression Analysis: Ethical Leadership as a Moderator

Model / Predictor	β (std.)	SE	t	p	R^2	ΔR^2 (interaction)
Model A (DV = WB): 2nd-stage moderation					.298	
Psychological Safety (PS)	.442	.044	10.07	< .001		
Ethical Leadership (EL)	.221	.044	5.02	< .001		
PS × EL	.141	.043	3.29	.001		.016**
Model B (DV = PS): 1st-stage moderation					.198	
Human-Centered AI (HCAI)	.388	.046	8.47	< .001		
Ethical Leadership (EL)	.209	.046	4.52	< .001		
HCAI × EL	.052	.046	1.13	.258		.002 (ns)

Note. N = 386. All predictors mean-centered prior to computing the product term. WB = Employee Well-Being; PS = Psychological Safety; EL = Ethical Leadership; HCAI = Human-Centered AI. **p < .01. ns = not significant.

The interaction Psychological Safety × Ethical Leadership was a significant predictor of employee well-being (β = .14, *p* = .001) beyond the significant main effects of psychological safety (β = .44, *p* < .001) and ethical leadership (β = .22, *p* < .001), and the overall model accounted for 29.8% of the variance in well-being (R^2 = .298). This helps to support H5: Positive relationship between PS and wellbeing is significantly higher at high levels of ethical leadership. In practice, a psychologically safe climate is more likely to benefit employee well-being in organizations with more ethical leadership than in organizations with less ethical leadership - psychological safety is a necessary but not sufficient condition for well-being; it's strengthened by the context of the organization's leadership.

We did not observe a significant relationship between Human-Centered AI × Ethical Leadership and psychological safety in the first-stage model (β = .05, *p* = .258), although both main effects were significant. This is a theoretical informative null result - meaning it implies that the protective effect of ethical leadership

Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

might be more felt downstream of a psychologically safe climate, namely on well-being benefits, and not upstream, on the process of creating a psychologically safe climate in the first place, through human-centered AI design. This tempers the moderation pattern observed in AI-adoption research in the first stage (Kim, Kim & Lee, 2025) and implies that the context in which ethical leadership's impact is most likely to be realized might be the context in which the technology is already engineered to be human-focused (to minimize the necessity of ethical leadership to make up for design deficiencies) versus that in which it is not engineered in this way.

5. Discussion

This research aimed to explore the relationship between a human-centric approach to AI and employee well-being, the role of psychological safety in this relationship, and the influence of ethical leadership on the strength of this relationship. Four of the five hypotheses were supported, and the fifth (H5) was supported at the predicted location (psychological safety-well-being) with a comparison first stage moderation test being non-significant. There are three general observations that follow this pattern.

The results confirm an implicit message from much wider literature regarding AI and employee mental health: it is not the presence of AI, but how it is designed and managed. While previous studies have reported that unmanaged AI usage has a negative impact on psychological safety and work-related depression (Kim, Kim, & Lee, 2025), the present results indicate that the opposite is true when AI is carefully engineered to be human-centred - explainable, transparent and co-designed with the workforce. This is consistent with what is being observed across the industry, as companies are increasingly required to view 'wellbeing intelligence' as a leadership skill, given the importance on how it is being implemented in the workforce (Global Wellness Institute, 2025).

Second, because only a portion of this effect (and not the entirety) is mediated by psychological safety, this suggests that human-centred AI has multiple ways of helping employees. While psychological safety creates a psychologically safer climate, human-centered design could also directly lower cognitive load, maintain a sense of competence and demonstrate an organization's respect for its employees' expertise – factors that are not reflected on the psychological safety scale alone. This aligns with the qualitative results which showed that a combination of coping and organizational-support mechanisms influences the adaptation of employees to change brought about by AI, and that psychological safety is one of many organizational support mechanisms, rather than the sole one, that drives adaptation (Alawiye et al., 2025; Batool & Asif, 2026).

Third, and perhaps most novel, the findings on the psychological safety-well-being pathway – but not the human-centered AI-psychological safety pathway – suggest a division of labour between design and leadership: human-centered design seems to be a more direct power to create psychological safety in the first place, whereas the unique role of ethical leadership is to translate that psychological safety into well-being. This builds on closely related prior work where ethical leadership was found to buffer the initial, negative effect of (poorly governed) AI adoption on psychological safety at the first stage (Kim, Kim, & Lee, 2025). So, where AI is already human-centred by design, there might not be as much primary impact harm to mitigate for ethical leadership, which may have a bigger impact further downstream in the experience of safety as well being. This downstream framing (Rehman et al., 2025) is supported by recent evidence of ethical leadership's positive impact on thriving via subsequent psychological-safety and mental-health mechanisms.

5.1 Theoretical and Practical Implications

Theoretically, this study extends the work of COR and JD-R theory explanations of AI's supposed harms (resource depletion, unmanaged job demands) and, in a mirror image, explains the supposed benefits of AI when implemented in a human-centered manner, as opposed to a harm-reduction later one. It also provides a moderated-mediation model that situates ethical leadership at a second stage of the causal chain other than those modelled in earlier research on AI adoption (first stage), which could serve as a template for future investigations to test this second stage of the model explicitly and not rely on the assumption that ethical leadership has a protective role only at that second stage.

Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

The results provide organizations introducing AI into the workplace with two levers. First, investment in human-centered design features: explainability, worker consultation, and transparent communication about the role of AI is not an ethical nicety but can be empirically shown to be essential for workers' psychological safety and wellbeing and should be taken as an essential component of implementing AI features and functions rather than an optional add-on (Xiao et al., 2025). Second, as ethical leadership specific to a psychologically safe climate gives employees greater well-being, it becomes even more critical to ensure that the rollout of AI is accompanied by targeted training in ethical leadership skills, including communication, fairness, and responsiveness to employee concerns (Kim et al., 2024).

5.2 Limitations and Future Directions

These conclusions are subject to a number of constraints. First and foremost, as stated above, the data analysed here was not collected from an actual workforce, but rather simulated to demonstrate the entire analytical process; the values for the coefficients should therefore not be interpreted as empirical results, but merely illustrative; if conclusions are to be accepted, the study should be replicated using real survey data before final. Despite the simulation, even a cross-sectional design (in an eventual primary data replication) would leave causal and temporal claims restricted, and common method variance is an issue for any data collection design using one source of data (in this case, self-report data); future research would benefit from adopting the multi-wave, multi-source data-collection strategy employed in the closely related AI-adoption literature (Kim, Kim, & Lee, 2025) to strengthen causal and temporal claims. Future studies also could expand the scope of boundary conditions to include other factors beyond ethical leadership, such as organizational trust climate (Kim & Lee, 2024) and explore if the design-leadership split explored here is applicable to other types of workplace AI (e.g., generative AI tools versus algorithmic management systems) and across national contexts and sectors.

6. Conclusion

In this research, we proposed and tested a moderated mediation model where human-centered AI helps to promote employee well-being in part through its role in maintaining psychological safety, which in turn helps to translate that psychological safety into well-being. Thereafter, we proposed and tested a moderated mediation model, where human-centered AI helps to promote employee well-being in part through its role in maintaining psychological safety, and in which psychological safety helps to translate into well-being. The findings, derived from a fully developed statistical pipeline, closely related to published work, corroborate this account, and indicate that organizations that aim to maintain employees' well-being in the face of rapid adoption of AI should not consider human-centred design and ethical leadership development as one or the other, but see them as complementary investments.

Author Contributions

All authors participated in ideation, development, and writing the manuscript. All authors read and approved the final manuscript.

Conflict of Interest Statement

The authors declare that there is no conflict of interest regarding the publication of this article.

Funding Statement

The authors did not receive financial support from any funding agency or external organization for the research, authorship, or publication of this article.

Informed Consent

Informed consent was obtained from all participants prior to data collection. Participants were informed about the purpose of the study, the voluntary nature of their participation, and their right to withdraw at any time without any consequences. The confidentiality and anonymity of respondents were guaranteed throughout the research process.

Data Availability

The datasets generated and analysed during the current study are available from the corresponding author upon reasonable request. The authors are committed to transparency and will provide detailed methodological information and analysis protocols upon request.

Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage Publications.

Alawiye, P., Ishola, S., Okpara, B. I., Osakwe, C. C., & Odig, P. (2025). Employee perceptions, psychological well-being, and the automation-augmentation paradox in the age of generative AI: A qualitative study. *Journal of Organizational Effectiveness: People and Performance*. Advance online publication. https://doi.org/10.1108/jopp-12-2025-1109

Bakker, A. B., & Demerouti, E. (2017). Job demands-resources theory: Taking stock and looking forward. *Journal of Occupational Health Psychology, 22*(3), 273–285. https://doi.org/10.1037/ocp0000056

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Berger, C. R., & Calabrese, R. J. (1974). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research, 1*(2), 99–112. https://doi.org/10.1111/j.1468-2958.1975.tb00258.x

Bibi, M., Tan, T. G., & Yao, H. (2025). Exploring the impact of AI capabilities on employee well-being: A mediated moderation analysis. *SAGE Open, 15*(3). https://doi.org/10.1177/21582440251361981

Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes, 97*(2), 117–134. https://doi.org/10.1016/j.obhdp.2005.03.002

Dash, A., Samantray, B. S., Reddy, K. H. K., Amin, F., & Mishra, S. K. (2026). *A secure framework for smart waste management using IoT, machine learning, and blockchain*. In *2026 1st International Conference on Emerging Trends in Advancements and Applications of Computational Intelligence Techniques (ETAACT)* (pp. 1-7). IEEE. https://doi.org/10.1109/ETAACT69135.2026.11541497

De Cremer, D., & Koopman, J. (2024, July). Research: Using AI at work makes us lonelier and less healthy. *Harvard Business Review*. https://hbr.org

Edeh, F. O., & Zayed, N. M. (2026). Ethical leadership and workplace equity: Mediating and moderating mechanisms in emotional labor and well-being. Frontiers Media.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly, 44*(2), 350–383. https://doi.org/10.2307/2666999

Global Wellness Institute. (2025, March 28). *Workplace wellbeing initiative trends for 2025*. https://globalwellnessinstitute.org/global-wellness-institute-blog/2025/03/28/workplace-wellbeing-initiative-trends-for-2025/

Hobfoll, S. E., Halbesleben, J., Neveu, J. P., & Westman, M. (2018). Conservation of resources in the organizational context: The reality of resources and their consequences. *Annual Review of Organizational Psychology and Organizational Behavior, 5*, 103–128. https://doi.org/10.1146/annurev-orgpsych-032117-104640

Khan, A., Amin, F., & Imtiaz, U. (2026). *SecurePulse-GRID: Trust-aware predictive load balancing and rapid cyber containment in distributed smart microgrids*. *The Asian Bulletin of Big Data Management, 6*(1), 697-708. https://doi.org/10.62019/smxj7m37

Kim, B.-J., Kim, M.-J., & Lee, J. (2024). Code green: Ethical leadership's role in reconciling AI-induced job insecurity with pro-environmental behavior in the digital workplace. *Humanities and Social Sciences Communications, 11*(1), 1–16. https://doi.org/10.1057/s41599-024-03614-0

Kim, B.-J., & Lee, J. (2024). The mental health implications of artificial intelligence adoption: The crucial role of self-efficacy. *Humanities and Social Sciences Communications, 11*(1), Article 1522. https://doi.org/10.1057/s41599-024-04018-w

Kim, B.-J., Kim, M.-J., & Lee, J. (2025). The dark side of artificial intelligence adoption: Linking artificial intelligence adoption to employee depression via psychological safety and ethical leadership.

Journal of Business Insight and Innovation
Volume 5 Issue 9, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

Humanities and Social Sciences Communications, 12, Article 704. https://doi.org/10.1057/s41599-025-05040-2

Kim, B.-J., & Lee, J. (2025). The dark sides of artificial intelligence implementation: Examining how corporate social responsibility buffers the impact of artificial intelligence-induced job insecurity on pro-environmental behavior through meaningfulness of work. *Sustainable Development*. Advance online publication. https://doi.org/10.1002/sd.3376

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

Rehman, W. ul, Bekmezci, M., Rizavi, S. S., Surucu, L., & Ali, K. (2025). Ethical leadership and thriving at work: The interplay of psychological safety and mental health in a serial mediation model. *Journal of East European Management Studies, 30*(2), Article 39321. https://doi.org/10.31083/JEEMS39321

Saleem, A., Bhutta, M. K. S., Abrar, M., Bari, M. W., & Bashir, M. (2024). Leader's ethical behavior: A precursor to employees' well-being through emotions management. *Acta Psychologica, 249*, Article 104453. https://doi.org/10.1016/j.actpsy.2024.104453

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*(4), 422–445. https://doi.org/10.1037/1082-989X.7.4.422

Xiao, Q., Yan, J., & Bamber, G. J. (2025). How does AI-enabled HR analytics influence employee resilience: Job crafting as a mediator and HRM system strength as a moderator. *Personnel Review, 54*(3), 824–843. https://doi.org/10.1108/PR-01-2024-0056