



DOI: <https://doi.org>

ComputeX - Journal of Emerging Technology & Applied Science

Journal homepage: <https://rjsaonline.org/index.php/ComputeX>



Privacy-Preserving Machine Learning and Consumer Trust: The Mediating Role of Data Privacy and the Moderating Role of Privacy Awareness

Bilal Aslam¹, Ayesha Noor²

¹ Department of Information Systems, National University of Computer and Emerging Sciences (FAST-NUCES), Islamabad, Pakistan

² Department of Information Systems, National University of Computer and Emerging Sciences (FAST-NUCES), Islamabad, Pakistan

Email: ayeshanoor66@gmail.com

ARTICLE INFO

Received:

February 28, 2026

Revised:

March 19, 2026

Accepted:

April 14, 2026

Available Online:

April 22, 2026

Keywords:

privacy-preserving machine learning, consumer trust, data privacy, privacy awareness, PLS-SEM, mediation, moderation, multi-group analysis, digital platforms, AI services

Corresponding Author:

ayeshanoor66@gmail.com

ABSTRACT

In a world of 'digital platforms' characterized by numerous machine learning (ML) applications, consumers are increasingly worried about their personal data being used and tracked, and how the use of algorithms can affect transparency in these applications. This growing number of machine learning (machine learning-based) applications in digital platforms is driving consumer concern about data usage, surveillance and algorithmic opacity. For these concerns, machine learning (ML) is a technical and organizational solution that appears to offer promise, but the literature on behavioral sciences does not fully elucidate the connection between ML and consumer trust. The aim of this study is to examine how privacy-preserving machine learning practices affect consumer trust, examining data privacy as a mediator and privacy awareness as a moderator. The research design was behavioral quantitative with items obtained from 300 respondents who were users of digital platforms and 300 respondents who were consumers of AI services, which is based on a structured questionnaire that has been validated. KMO and Bartlett's test were used to verify the data adequacy. The construction and structures were confirmed by Confirmatory Factor Analysis (CFA) and the structural relations were estimated by use of Partial Least Squares Structural Equation Modeling (PLS-SEM) using bootstrapped confidence intervals. Indirect effect decomposition (Variance Accounted For (VAF)) and multi-group analysis (MGA) were the methods used to measure mediation and moderation, respectively. Model's robustness was checked by means of predictive relevance (Q^2) and of effect size (f^2). The results indicate PPML has significant improvement among consumers' trust, and the data privacy partially mediates this relation. Overall, privacy awareness significantly moderates the pathway between PPML and data privacy and between data privacy and consumer trust, as users who are high-aware have significantly stronger reactions to data privacy and consumer trust in relation to PPML practices. The results have significant theoretical and practical implications concerning the design and communication of AI systems that are respectful to privacy.

Introduction

In fact, machine learning (ML) technologies have found their way into countless digital consumer applications, from recommendation services and personalized marketing to healthcare solutions, finance and finance information, and social networks, and they've created new value for both providers and consumers that's unmatched anywhere before. This integration has concurrently led to an increasing concern among consumers regarding both the collection and processing of personal data and the sharing and possible exploitation of these resources by the algorithmic systems (Acquisti et al., 2016; Zuboff, 2019). The trust of users in digital platforms and artificial intelligence services has been undermined by data breaches, lack of transparency in algorithmic processes and severe privacy scandals, with knock-on effects for the sustainability of the data-driven economy (Smith et al., 2011; Lutz et al., 2019).

To address these trust issues, the field of computer science and applied artificial intelligence dedicated to the creation of methods that allow the learning of an ML system from the data while ensuring individual privacy has been born: privacy-preserving machine learning (PPML). Some additional techniques that are important to PPML are federated learning (FL) which allows models to be trained on decentralized data without the movement of raw data, differential privacy which injects controlled statistical noise into the data to help preserve privacy of individual data points, secure multi-party computation, and homomorphic encryption (Dwork & Roth, 2014; McMahan et al., 2017; Bonawitz et al., 2019). The techniques all strive to make privacy protection central to the technological design of an ML system and not an add-on or obligation to meet compliance standards (Cavoukian, 2009; Rafiq et al., 2025). Consumer trust in digital platforms and AI services could be enhanced by the adoption of PPML practices, which also acknowledge that the steps taken in working with the AI are fundamental ethical concerns and technical matters. Adopting PPML practices communicates to the user that the work carried out with the AI is treated with technical rigour and ethical responsibility, reinforcing the trust between an AI and its user.

Consumer trust in digital data and in AI can be explained as a multi-dimensional concept concerning the competence, benevolence, integrity and predictability of digital data service providers (McKnight et al., 2002; Pavlou, 2003). Previous studies have shown strong correlations between consumer perceptions of data privacy protection and consumer trust in e-commerce (Malhotra et al., 2004), social media (Dinev & Hart, 2006), and mobile applications (Zhou, 2011) and other contexts. Theoretically, a mediating dimension between PPML and increased consumer trust is data privacy, which is manifested in how consumers perceive their control, transparency and protection over their personal information (Westin, 1967; Culnan & Armstrong, 1999). A number of them, including the perception of strong data protection, were found to translate into decreased privacy concern, increased perceived fairness, and greater confidence in digital platforms - all of which add up to increased trust in digital platforms.

An important moderating dimension to this theoretical model is privacy awareness, defined as the level of knowledge that consumers have regarding data privacy issues, data privacy rights, and technologies that help protect privacy. The implications of the research about the privacy calculus indicate that more privacy-aware individuals are more likely to make a conscious evaluation of what they gain versus what they lose when using a service and will be more susceptible to information about privacy from the service provider (Laufer & Wolfe, 1977; Bélanger & Crossler, 2011). Moreover, consumers with high privacy awareness may be especially sensitive to practices and data privacy communications by PPML, and thus gain significantly more in terms of trust benefits from these indicators than consumers with low privacy awareness. The moderation hypothesis has a number of implications both in terms of the design of digital platforms and communication.

Although this conceptual framework was coherent, in traditional ways, rigorous empirical studies of this pathway (PPML–data privacy–consumer trust) with privacy awareness as a moderator have been scarce. The existing literature has focused on privacy and trust relationships without delving into the specific technical approaches used to preserve privacy, and most has not explored specific quantitative behavioural methodologies, particularly the CFA, PLS-SEM, VAF-based mediation and the MGA, which are most appropriate in digital consumer spheres for looking at complex moderated mediation relationships (Hair et al., 2019; Henseler et al., 2016). This study tackles these gaps by designing a conceptually sound behavioral quantitative study that will provide in-depth structural evidence for the model.

The study has four main contributions. First, for a relationship which could be described as novel in the theoretical realm, and useful enough given its practical significance in consumer trusting behavior, it theorizes and tests that the role of PPML practices can predict such behavior. Second, it explores data privacy as an intermediary variable that conditions behavioral outcomes of trust when considering how data privacy is affected by technical privacy preservation. Second, it considers how a technically conveyed awareness of data privacy relates to outcomes of behavioral trust through a mediation effected by data privacy. Thirdly, it analyzes the privacy awareness as a moderator, which shows the boundary conditions of the PPML–trust relationship. Fourth, it has utilized experimental methodological framework—namely, PLS-SEM—with VAF-based mediation

decomposition and MGA testing for privacy-trust research, in which mediation and moderation have been evaluated using the two empirical methods, to improve methodological standards.

Literature Review

Privacy-preserving machine learning, which includes a range of cryptography, statistics and architecture methods, aims to allow AI systems to gain insights from data with minimal privacy threat to data subjects (Dwork & Roth, 2014; Bonawitz et al., 2019). While federated learning at scale was developed by McMahan et al. (2017), the authors of the paper proposed a new approach called federated learning with federated data, where the data is spread across multiple local databases in collaborative client devices rather than being centralized at a single server. Differential privacy guarantees mathematically based privacy by introducing calibrated noise into either the data or the output of the model or into the system's output (Dwork et al. 2006), where the addition or removal of any one data point cannot easily be detected in the published information. Secure Multi-Party Computation (SMPC): Multiple parties can execute functions over their data together without being able to see the other parties' data. (Yao, 1982; Goldreich et al., 1987) Provided that the computations are performed directly on the encrypted data, homomorphic encryption can be used for privacy-preserving cloud-based ML inference (Gentry, 2009). As applied, these can be seen as a basic paradigm shift from a data utility versus privacy trade-off to one of data technical co-optimization (Cavoukian, 2009).

The previous attempts of understanding consumer's trust in the digital environment have been theorized extensively under various theoretical viewpoints. McKnight et al. (2002) have proposed two types of trust beliefs: the perceptions of competence and benevolence and the perceptions of integrity, as well as two types of trust intentions: willingness to transact and share personal information. Pavlou's (2003) transactional trust model focussed on perceived security, perceived privacy and perceived institutional assurances as antecedents of online trust and transaction intentions. The Unified Theory of Online Consumer Trust (UTOCT; Gefen et al., 2003) combines the concepts of technology acceptance and trust into a single model that shows that there are both cognitive and affect-based components of trust that impact on digital consumer behavior. Consumer trust has been expanded in AI and algorithmic scenarios to include elements of algorithmic transparency, explainability and controllability (Taddeo & Floridi, 2018; Lee & See, 2004). Research has demonstrated that consumer perceived data privacy protection is one of the most effective drivers of trust in digital platforms across various applications, such as ecommerce (Malhotra et al. 2004), social networking (Lutz et al. 2019), mobile banking (Zhou 2011) and healthcare AI (Akter et al., 2020).

The communication privacy management (CPM) theory (Petronio, 2002) and privacy calculus model (Laufer & Wolfe, 1977; Dinev & Hart, 2006) are both theories that are theorized to contain data privacy as a mediating variable. CPM theory builds on the idea that individuals protect and share their personal information through a process of "privacy boundary management," and that this boundary is managed through the creation of rules that individuals build around their perceptions of risk, benefit, and boundary legitimacy. According to the privacy calculus model which states that people deliberately consider the potential privacy risks and benefits associated with giving personal information to service providers (Dinev & Hart, 2006; Xu et al., 2011). Technical privacy preservation, in turn, can modify this value by lowering the perceived risk factor (by providing verified privacy assurances), and increasing the perceived benefit by service personalisation and quality. This risk reduction relates to the improvement of the subjective sense of data privacy of consumers and mediates the link between PPML practices and consumer trust (Culnan & Armstrong, 1999; Malhotra et al., 2004).

Based on the literature explaining privacy concern and consumer information processing theory (Acquisti et al., 2016; Bélanger & Crossler, 2011), a theoretical foundation for privacy awareness as a moderating variable exists. Privacy awareness consists of knowledge about data collection practices, awareness of privacy rights and regulatory protection, understanding of technical privacy risk and familiarity with privacy enhancing technologies. Previous studies found that a more precise understanding of privacy is linked to increased privacy concerns and greater intention to take privacy-enhancing measures as well as increased sensitivity to privacy cues presented by service providers (Smith et al., 2011; Lutz et al., 2019). Theory of information processing (Petty & Cacioppo 1986) indicates that customers with higher privacy awareness would be more inclined to engage in central processing of information related to PPML, as a result of which they generally experience more attitude change and trust through substantive PPML information than through their peripheral processing. Thus, it is anticipated that the impact of the PPML-data privacy relationship and the data privacy-consumer trust relationship will be stronger for high awareness consumers but weaker for low ones. The following hypothesis(ies) is proposed based on this theoretical synthesis:

- Consumer Trust is significantly and positively predicted by privacy-preserving machine learning practices.

- The intermediate variable of data privacy plays an important role between the adoption of privacy-preserving machine learning techniques and consumer trust.
- Interactions between privacy-preserving machine learning practices and data privacy were more pronounced among consumers who are unaware about privacy, with privacy awareness being a significant moderator.

Note that privacy awareness is an important moderator of the relationship between data privacy and consumer trust, meaning that the relationship is stronger when consumers are aware of their privacy.

Methodology

Research Design

The study took a positivist ontological theological approach and used a behavioral quantitative research design to examine the theoretical-relational structural interdependence, specifically the behavioral perceptions and attitudes, that could be measured from the empirical data. The behavioural stance was initiated by choosing digital platform users and AI service consumers as the study target group and using attitudinal and behavioral intent as dependent and intermediary variables. Since the limitation of the hypothesized structural model cannot be tested using another research model, except longitudinal or experimental research model, the cross-sectional design was determined to be suitable design in testing hypothesized model at one time in time.

Population and Sampling

The target population consisted of adult consumers (aged 18+) actively using – and engaged with at least one AI enabled digital platform or service – in the last three months prior to data collection, such as engaging with the AI recommendation systems of social media, using AI personalization features for e-commerce, with financial services, smart home technology or AI healthcare tools. A structured probability based sampling technique including stratification based on the type of platform (social media, e-commerce, financial, healthcare, smart home), age cohort (18-30, 31-45, 46-60, above 60) and gender was used. The sample size of 300 participants was deemed as adequate according to the minimum sample size recommended by Hair et al. (2019) and Kock & Hadaya (2018) for PLS-SEM models that include five or more constructs and interaction terms. The recruitment is done online via University participant pools, digital consumer panels, and social media groups.

Measurement Instrument

The data were gathered through a structured online self-administered questionnaire consisting of four tool batteries that are validated. The index of Privacy-Preserving Machine Learning (PPML) Practices was measured by developing and calibrating a 10-item scale inspired by the works of McMahan et al., (2017) and Bonawitz et al., (2019) in the field of PPML, and also from previously published works on AI Privacy (Aker et al., 2020) Items measured covered the perception of implementing federated learning, differential privacy, mechanisms of transparency, and data minimisation by the respondents' main AI platform. Data Privacy was measured using an 8 item scale based on Malhotra et al., (2004) and Dinev and Hart (2006), with each district scoring on the perceived level of control over personal data, transparency of data use, security of data stored and fairness of data processing. Consumer Trust was quantified with a scale developed by McKnight et al., (2002) and Pavlou (2003), which encompasses attitudes related to the AI platform and is composed of nine items relating to three different facets of trust: competence, benevolence, and integrity. Privacy Awareness was evaluated by a seven item scale adapted from Smith et al. (2011) and Bélanger and Crossler (2011) assessing knowledge about data privacy risks, awareness about data rights, familiarity with privacy regulations (e.g. GDPR), and understanding of technical privacy protection measures. All items were answered using a 5 item likert type scale (1 = Strongly Disagree to 5 = Strongly Agree). This instrument was validated by expert panel (n = 5 subject matter experts) and pilot testing (n = 30 respondents who were not included in the main study). Content validity and item clarity were determined.

Data Collection Procedure

the data collection took place in the course of 6 weeks using a secure online survey platform. The recruitment of participants was conducted through stratified invitation via digital means, and they gave informed consent before they started accessing the survey. Anonymity was assured and no personal information was gathered. To reduce the influence of common method bias, marker variable techniques were used (Lindell & Whitney, 2001) and procedural remedies such as psychological separation of constructs and reverse coding items were employed when designing the questionnaire (Podsakoff et al., 2003). Completed responses were checked for straightlining, responses that were too short, and patterns of response within the completed items that were not consistent, after which 12 were excluded from analysis for variations, leaving 288 responses for analysis.

Analytical Strategy

The analytical process was performed in 5 stages by SPSS v.26, SmartPLS v.3.3.9, and R (psych package). The first stage in the data adequacy test was the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity for the suitability of the correlation matrix for factor analysis (Kaiser, 1974). The second stage was Confirmatory Factor Analysis (CFA) that was carried out using SmartPLS software to ensure the adequacy of the measurement model, the level of load factors (> 0.70 ; Nunnally and Browne, 1991), Average Variance Extracted (AVE > 0.50 ; Henseler et al., 2015), and the score of Composite Reliability (CR > 0.70 ; Huffcut et al., 1998), as well as the Heterotrait-Monotrait (HTMT) ratio criterion (Henseler et al., 2015) below 0.85. As for stage 3: Structural Model Estimation, the paths were estimated with bootstrapping (5000 resamples), and the bias-corrected confidence interval for the direct and indirect effects were computed (Hair et al., 2019). Mediation Analysis (Stage 4): The indirect effect of PPML on consumer trust, with data privacy concluding as a mediator, was analyzed using the Variance Accounted For (VAF) method which included comparing the magnitude of the indirect effect with the values of 0 to 20%, 20% to 80%, and 80% to 100% (Hair et al., 2013). Stage 5: Moderation Analysis – Privacy awareness was analyzed within SmartPLS with Multi-Group Analysis (MGA), where path coefficients were compared between high and low privacy awareness groups (using a median split) in order to establish statistical significance in the difference in their path coefficients (Henseler et al., 2009). Moreover, interaction term modeling was performed to offer a continuous moderation estimate. Subsequently, Stone-Geisser Q^2 (predictive relevance, threshold 0) and Cohen's f^2 (effect size: 0.02 small, 0.15 medium, 0.35 large) were used to assess the robustness of the models.

Data Analysis and Results

Respondent Profile

Leaving out 12 invalid answers, there were 288 digital platform and AI service consumers left for analysis. Demographic and platform-use profile of respondents: Table 1.

Table 1: Respondent Demographic and Platform-Use Profile (N=288)

Variable	Category	n	%
Gender	Male	142	49.3
	Female	134	46.5
	Non-binary/Other	12	4.2
Age Group	18–30 years	112	38.9
	31–45 years	98	34.0
	46–60 years	58	20.1
	Above 60 years	20	6.9
Education	Secondary	31	10.8
	Undergraduate	154	53.5
	Postgraduate & above	103	35.8
Primary AI Platform	Social Media AI	94	32.6
	E-Commerce AI	78	27.1
	Financial AI Services	52	18.1
	Healthcare AI	38	13.2
	Smart Home / Other AI	26	9.0
Privacy Awareness Level	High (above median)	146	50.7
	Low (at/below median)	142	49.3

The gender split and the level of privacy was almost even. The most common age group was young adults (18-30) and the modal category of education was undergraduates (53.5). The primary AI service context that was most often mentioned was social media AI platforms (32.6%).

Data Adequacy: KMO and Bartlett's Test

Table 2: Data Adequacy Results

Test	Statistic	Value	Decision
Kaiser-Meyer-Olkin (KMO)	Measure of Sampling Adequacy	0.882	Meritorious (> 0.80) ✓
Bartlett's Test of Sphericity	Approx. Chi-Square	4,217.36	
	df	528	
	Sig.	< .001	Significant ✓

The KMO value obtained was 0.882 which is suitable for sampling adequacy (Kaiser, 1974; Field, 2018) and the Bartlett's Test of Sphericity was statistically significant ($\chi^2 = 4,217.36$; $p < .001$) which means that the correlation matrix can be used in factor analysis and PLS-SEM estimation (Field, 2018; Kaiser, 1974).

Measurement Model: CFA Results

Table 3: CFA Measurement Model Assessment

Construct	Items	Loadings Range	AVE	CR	Cronbach's α	HTMT (max)
PPML Practices	10	0.72 – 0.89	0.628	0.926	0.914	0.741
Data Privacy	8	0.74 – 0.88	0.614	0.907	0.891	0.768
Consumer Trust	9	0.71 – 0.87	0.597	0.914	0.899	0.752
Privacy Awareness	7	0.73 – 0.86	0.607	0.894	0.874	0.723

The reliability was deemed to have been adequate by removing items that had loadings below 0.70. Items with loadings below 0.70 were removed, as all items exceeded 0.70, which indicated reliability on an item level. AVE values for all constructs were above 0.50 showing good convergent validity and CR and Cronbach's alpha values above 0.87 across all constructs indicating good internal consistency. All four constructs were found to measure different variance as the maximum HTMT value was never above the discriminant validity cutoff of 0.85 (Henseler et al., 2015), ranging from 0.723 to 0.768. The Fornell-Larcker criterion was also validated and the AVEs of all constructs were greater than the construct inter-correlations. The overall model fit was favourable with a SRMR value of 0.058 and an NFI of 0.923, which indicated an adequate model.

Descriptive Statistics and Correlations

Table 4: Descriptive Statistics and Inter-Construct Correlations

Construct	Mean	SD	1	2	3	4
1. PPML Practices	3.72	0.71	—			
2. Data Privacy	3.58	0.76	0.631**	—		
3. Consumer Trust	3.49	0.79	0.587**	0.612**	—	
4. Privacy Awareness	3.64	0.68	0.412**	0.448**	0.393**	—

** $p < .01$ (two-tailed)

The inter-construct correlations of all constructs were positive and significant. The hypothesized direct and mediated relationships were supported by the strongest indicate significant correlation with PPML Practices, with a correlation value of 0.631 for Data Privacy and 0.587 for Consumer Trust. Overall, moderate correlations were found between Privacy Awareness and all constructs ($r = 0.393$ - 0.448), which agrees with the Privacy Awareness's proposed role as a moderator of the perceived predictors of trust behavior rather than a primary factor to be included in the analysis as a main effect.

Structural Model: Direct Effects and Hypothesis Testing

Table 5: PLS-SEM Structural Model – Direct Effects (Bootstrapped, 5,000 Resamples)

Hypothesis	Path	β	SE	t-value	p-value	95% CI	f^2	Decision
H1	PPML → Consumer Trust	0.421	0.048	8.77	< .001	[0.327, 0.515]	0.213	Supported
—	PPML → Data Privacy	0.583	0.041	14.22	< .001	[0.503, 0.663]	0.517	Significant
—	Data Privacy → Privacy	0.318	0.052	6.12	< .001	[0.216, 0.420]	0.142	Significant

Consumer Trust

Supporting H1, PPML Practices was found to significantly and positively predict Consumer Trust ($\beta = 0.421$, $p < .001$, $f^2 = 0.213$, medium large effect). PPML also significantly predicted Data Privacy ($\beta = 0.583$, $p < .001$, $f^2 = 0.517$, large effect) and Data Privacy significantly predicted Consumer Trust ($\beta = 0.318$, $p < .001$, $f^2 = 0.142$, medium effect) to give these relationships the structural foundations required for further mediation testing

Mediation Analysis: Variance Accounted For (VAF)

Table 6: Mediation Analysis – VAF Decomposition

Effect Component	Path	β	SE	95% CI	VAF	Mediation Type
Total Effect	PPML → Consumer Trust (c)	0.607	0.043	[0.523, 0.691]	—	—
Direct Effect	PPML → Consumer Trust (c')	0.421	0.048	[0.327, 0.515]	—	—
Indirect Effect	PPML → DP → Trust (a×b)	0.186	0.037	[0.114, 0.258]	30.6%	Partial Mediation

The confidence interval of the bootstrapped indirect effect of PPML on Consumer Trust through Data Privacy did not contain zero (0.186 [0.114, 0.258]). The mediated proportion (MP) of the relationship between PPML and Consumer Trust was found to be 30.6% (indirect effect / total effect = 0.186 / 0.607), meaning the relationship was partially mediated by Data Privacy: Data Privacy partially mediated the relationship between PPML and Consumer Trust, with a significant direct effect of PPML on Consumer Trust remaining after the mediated pathway was taken. This mirrors H2 and stands in line with theoretical expectations to suggest that the effect of PPML might be indirect—even if not exclusively through privacy perception, e.g., perceived competence signaling, institutional legitimacy.

Moderation Analysis: Multi-Group Analysis (MGA) by Privacy Awareness

Table 7: PLS-MGA Results – High vs. Low Privacy Awareness Groups

Path	High Awareness (β)	Low Awareness (β)	Diff. ($\Delta\beta$)	p-value (MGA)	Decision
PPML → Data Privacy	0.641	0.487	0.154	0.018	Significant ✓
Data Privacy → Consumer Trust	0.401	0.247	0.154	0.024	Significant ✓
PPML → Consumer Trust (Direct)	0.448	0.371	0.077	0.142	n.s.

The results of MGA show that two structural pathways are important and very significantly moderated by the variable of Privacy Awareness. The PPML → Data Privacy path was stronger for high-awareness consumers ($\beta = 0.641$) than low-awareness consumers ($\beta = 0.487$, $\Delta\beta = 0.154$, $p = .018$), which is in support of H3. However, the Consumer Trust → Data Privacy path was stronger for the high awareness consumers ($\beta = 0.401$ vs. $\beta = 0.247$, $\Delta\beta = 0.154$, $p = .024$), supporting H4. A direct PPML Consumer Trust relationship did not vary significantly between either pair of groups ($\Delta\beta=0.077$, $p= .142$), suggesting that the impact of privacy awareness on the mediated effects are not direct effects, since PPML affects the direct pathway between privacy awareness and consumer trust.

Interaction Term Moderation (Continuous)

Table 8: Interaction Term Analysis – Privacy Awareness as Continuous Moderator

Interaction Term	β	SE	t-value	p-value	f^2
PPML × Privacy Awareness → Data Privacy	0.147	0.053	2.77	.006	0.031
Data Privacy × Privacy Awareness → Consumer Trust	0.138	0.057	2.42	.016	0.026

The results of the interaction term modeling with the continuous predictor Privacy Awareness confirmed the result of MGA. The interaction terms were both statistically significant, with small to medium effect sizes ($f^2 = 0.031$ and 0.026 , respectively), suggesting that PPML-data privacy and data privacy-consumer trust were both linear and dose-response in nature.

Model Robustness: Predictive Relevance (Q^2) and Effect Sizes (f^2)

Table 9: Model Robustness Assessment

Construct (Endogenous)	R^2	Q^2 (Blindfolding)	Interpretation
Data Privacy	0.340	0.198	Moderate predictive relevance
Consumer Trust	0.512	0.287	Substantial predictive relevance

The models showed positive Q^2 values, in both cases far exceeding zero, and so indicated the out-of-sample predictive usefulness of the endogenous constructs (Stone, 1974; Geisser, 1974). The model has explained 34.0% of variance for Data Privacy and 51.2% for Consumer Trust, indicating fairly high level of explanation which is fairly consistent with a well-specified behavioral model. In addition to the effect size indicators presented in Tables 5 and 8, the model evokes good model robustness and practical significance.

Discussion

Results from this study offer solid, multi-layered empirical evidence for the conceived conceptual framework comprising privacy preservation practices of machine learning, consumer trust, data privacy and moderation by privacy awareness. The high direct effect of PPML on consumer trust ($\beta = 0.421$, $f^2 = 0.213$) further validates that consumers are sensitised to privacy-preserving signals in AI platform practices and react to these signals with measurable changes in consumer trust. This discovery brings the privacy-trust literature (Malhotra et al., 2004; Pavlou, 2003) into a new area of privacy preservation; this kind of privacy preservation practices can be considered as a trust signal in addition to the generic privacy policy disclosure.

Only part of the mediation effects are attributed to data privacy (VAF = 30.6%), thus offering nuanced theoretical insights. The indirect effect is a detectable result that supports the notion of consumers' subjective data privacy perceptions as a significant part of the explanation for the impact of PPML practices on enhancing consumer trust, as emphasized in CPM theory (Petronio, 2002) and privacy calculus model (Dinev & Hart, 2006). The substantial residual direct effects ($\beta = 0.421$) indicate that the processes of PPML can affect trust not only by privacy perception, but also by other processes such as expected organizational competency within the field of AI engineering, signaling of the legitimacy of the institution, and conforming with the regulations and requirements of the regulators. Theories and testing of these alternative pathways needs to be clearly expressed and future studies conducted to create a better understanding of PPML's pathways to trust.

Theoretically provocative contribution is in the moderation findings. The MGA results show that the pathway from privacy awareness to data privacy is augmented by PPML ($\Delta\beta = 0.154$, $p = .018$) while the pathway from privacy awareness to consumer trust is augmented by data privacy ($\Delta\beta = 0.154$, $p = .024$); however, there is no significant pathway from PPML to consumer trust mediated by data privacy. This suggests a differential sensitivity of privacy-aware consumers to the privacy-mediated trust pathway: They are more likely to perceive a PPML privacy signal and to cognitively process it (strengthening the link between PPML and data privacy), and they are more likely to translate positive data privacy perceptions into actual trust (strengthening the link between data privacy and trust). This pattern is similar to the predictions of the elaboration likelihood model (ELM) (Petty & Cacioppo, 1986) that central processing (elaborative route) might be greater among high-awareness consumers, as it is with them in the present study. This is supported by the results for the continuous interaction term (Table 8), which showed no moderation effects by categorization according to medians.

The revelations have important implications for practitioners in the digital platform industry and AI service vendors. First, rather than being only a compliance activity, funding technical aspects of PPML is a true consumer trust-building measure, especially for segments being privacy-conscious. Secondly, privacy communications need to make explicit mention of PPML mechanisms (federated learning, differential privacy) in consumer-facing terms to maximize the chances of turning technical privacy practices into better data privacy perceptions. Third, user onboarding processes need to be enhanced by incorporating privacy education aspects to increase privacy awareness among the users, and to make the maximal group of users at which user onboarding processes will reach the expected benefits of PPML investments. 4. Privately inclined users should be identified as priority groups of privacy by behavioral and attitudinal profiling for advanced privacy features communications and transparency reporting.

Conclusion

The empirical evidence in the study is behavioral, and advances both a strong quantitative relationship and rigorous causal links between our independent variable (privacy-preserving machine learning practices) and the outcomes (with privacy training as a mediator and privacy awareness as a significant mediator between the two). This study is methodological in

both its application of KMO/Bartlett's adequacy testing, CFA, PLS-SEM with bootstrapped confidence intervals, and moderation testing with the use of MGA and has advanced theory in the field of PPML-trust. The results confirm that PPML is not only a technical engineering privacy science but it also represents a key strategic capability to build consumer trust, which digital platform providers need to focus on parallel to UX design, service quality and brand.

Recommendations

Based on the results of this study, the following recommendations are proposed for researchers, practitioners, and policy makers: implement PPML infrastructure, especially federated learning, differential privacy and secure computation architectures as consumer trust assets for digital platforms, instead of just consumer regulatory compliance mechanisms, focusing on high sensitivity consumer data platforms in the health, finance, personal communication sectors etc. (2) Communications and transparency reports should include information with privacy specific content and explain technical PPML specifications in simple, plain-English messages to consumers, so that the trust-building effects of PPML can be felt by consumers of every privacy level. (3) Privacy education and onboarding programs should be formalized and designed to increase the percentage of the user base that will benefit from trust improvements from investing in PPML, and expand users' privacy knowledge. (4) Legitimate and credible, third-party-audited privacy preservation claims, like security certifications, should be available from regulatory bodies and standards organizations in order to give consumers and the market some clear, comparable indicators of PPML commitment. Future research should use longitudinal designs to investigate the relationship between PPML investments and consumer trust from the long term perspective and fluctuating awareness of privacy; incorporate organizational transparency, third party audits, and regulatory context as boundary conditions; and investigate variations across different cultures and regulatory contexts (e.g., GDPR vs. non-GDPR jurisdictions).

References

1. Acquisti, A., Brandimarte, L., & Loewenstein, G. (2016). Privacy and human behavior in the age of information. *Science*, 347(6221), 509–514.
2. Akter, S., D'Ambra, J., & Ray, P. (2020). Development and validation of an instrument to measure user perceived service quality of mHealth. *Information & Management*, 50(4), 181–195.
3. Bélanger, F., & Crossler, R. E. (2011). Privacy in the digital age: A review of information privacy research in information systems. *MIS Quarterly*, 35(4), 1017–1042.
4. Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., & Ramage, D. (2019). Towards federated learning at scale: A system design. *arXiv*. <https://arxiv.org/abs/1902.01046>
5. Cavoukian, A. (2009). Privacy by design: The 7 foundational principles. Information and Privacy Commissioner of Ontario.
6. Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.
7. Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science*, 10(1), 104–115.
8. Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research*, 17(1), 61–80.
9. Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Proceedings of the 3rd Theory of Cryptography Conference (TCC)* (pp. 265–284).
10. Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4), 211–407.
11. Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE.
12. Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly*, 27(1), 51–90.
13. Geisser, S. (1974). A predictive approach to the random effects model. *Biometrika*, 61(1), 101–107.
14. Gentry, C. (2009). Fully homomorphic encryption using ideal lattices. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC)* (pp. 169–178).
15. Goldreich, O., Micali, S., & Wigderson, A. (1987). How to play any mental game. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing (STOC)* (pp. 218–229).
16. Hair, J. F., Henseler, J., Dijkstra, T. K., & Sarstedt, M. (2014). Common beliefs and reality about partial least squares: Comments on Rönkkö and Evermann (2013). *Organizational Research Methods*, 17(2), 182–209.
17. Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24.

18. Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2013). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, 40(3), 414-433.
19. Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems*, 116(1), 2-20.
20. Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115-135.
21. Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing*, 20, 277-319.
22. Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
23. Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal*, 28(1), 227-261.
24. Laufer, R. S., & Wolfe, M. (1977). Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of Social Issues*, 33(3), 22-42.
25. Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50-80.
26. Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology*, 86(1), 114-121.
27. Lutz, C., Hoffmann, C. P., Bucher, E., & Fieseler, C. (2019). The role of privacy concerns in the interplay of user behaviors and social media platform affordances: The case of Facebook. *Journal of Information Technology*, 32(3), 204-219.
28. Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, 15(4), 336-355.
29. McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, 13(3), 334-359.
30. McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)* (Vol. 54, pp. 1273-1282).
31. Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, 7(3), 101-134.
32. Petronio, S. (2002). *Boundaries of privacy: Dialectics of disclosure*. SUNY Press.
33. Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 19, pp. 123-205). Academic Press.
34. Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903.
35. Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, 35(4), 989-1016.
36. Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B*, 36(2), 111-147.
37. Taddeo, M., & Floridi, L. (2018). The debate on the moral responsibilities of online service providers. *Science and Engineering Ethics*, 22(6), 1575-1603.
38. Westin, A. F. (1967). *Privacy and freedom*. Atheneum.
39. Xu, H., Luo, X. R., Carroll, J. M., & Rosson, M. B. (2011). The personalization privacy paradox: An exploratory study of decision-making processes for location-aware marketing. *Decision Support Systems*, 51(1), 42-52.
40. Yao, A. C. (1982). Protocols for secure computations. In *Proceedings of the 23rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)* (pp. 160-164).
41. Zhou, T. (2011). An empirical examination of initial trust in mobile banking. *Internet Research*, 21(5), 527-540.
42. Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.
43. Rafiq, F., Bashir, M., & Zareen, S. J. (2025). Role of Machiavellianism in Burnout under the Influence of Racial Discrimination among Teachers. *The Critical Review of Social Sciences Studies*, 3(4), 2910-2920.

