International Multidisciplinary Journal of Emerging Technologies and Applications (IMJETA)
Vol. 1, No. 1, pp. 30-43, April 2026
Received April 04, 2026; Revised April 05, 2026; Accepted April 07, 2026
Published April 16, 2026

Presence, Embodiment, and Emotional Response in Mixed Reality Product Prototyping: Toward a Psychophysiological Model for XR-Based Design Evaluation

Septien Dwi Savandha[1][0009-0006-0106-3398]*, Adelia Azzahra[2][0009-0006-2624-112X], Elvira Fitriyanti[3][0009-0003-9547-3248], Nabilaa Faizatuz Zuhriyah[4][0000-0003-0391-3588], Pegi Sugiartini[5][0009-0008-2231-8833]
[1]Universidad Tecnológica Latinoamericana en Línea (UTEL), Mexico
dwisavandha9@gmail.com*
[2]Swadaya Gunung Jati University, Indonesia
adeliaazzahra349@gmail.com
[3]International Polytechnic Siber Cerdika, Indonesia
elvirafitriyanti@gmail.com
[4]Islam Bunga Bangsa Cirebon University, Indonesia
nabilaafaizatuzzuhriyah13@gmail.com
[5]Muhammadiyah Cirebon University, Indonesia
pegisugiartini@gmail.com

Abstract. The integration of extended reality technologies into industrial product design evaluation introduces a class of psychological variables, such as sense of presence, embodied cognition, and emotional arousal, that are absent from conventional flat-screen computer-aided design environments. Despite growing adoption of immersive head-mounted display systems for prototype assessment, no empirically validated model exists to explain how these psychophysiological states influence the quality and reliability of design judgments produced within immersive conditions. This study proposes and tests a four-layer psychophysiological framework linking extended reality modality type to design judgment reliability through three simultaneous mediators: subjective presence, galvanic skin response amplitude, and eye-tracking fixation duration. A within-subject experiment was conducted with 48 industrial designers and mechanical engineers who evaluated an identical computer-aided design prototype across four counterbalanced conditions: flat-screen display, augmented reality, and mixed reality. Heart rate variability, galvanic skin response, and fixation duration were recorded continuously alongside the ITC-Sense of Presence Inventory and a semantic differential judgment scale. Structural equation modelling confirmed full mediation of the display-condition effect on judgment reliability by the three psychophysiological pathways, with emotional arousal emerging as the strongest mediator. Extended reality experience level moderated arousal responses in expert users, preserving judgment reliability across

immersive conditions. These findings offer the first empirically grounded framework linking immersive display technology to evaluative cognition in engineering design contexts.

Keywords: Extended Reality; Design Evaluation; Psychophysiological Response; Sense of Presence; XR-Based Design Evaluation.

1. Introduction

The rapid proliferation of extended reality (XR) technologies, including virtual reality (VR), augmented reality (AR), and mixed reality (MR), has fundamentally transformed the methods industrial designers and engineers use to evaluate product prototypes (Kamińska et al., 2022; Larsen et al., 2024). Traditionally, prototype evaluation has relied upon flat-screen computer-aided design (CAD) environments in which the evaluator occupies a detached, observer-oriented posture toward the design artefact. Within this paradigm, judgments regarding form, ergonomics, and functional appropriateness are made under conditions of minimal sensory engagement and low emotional activation (Zhang et al., 2024). The advent of head-mounted display (HMD) systems capable of generating high-fidelity, immersive environments has disrupted this paradigm, enabling design teams to inhabit and interact with full-scale three-dimensional representations of products prior to physical fabrication. This transition carries substantial promise for accelerating design iteration cycles and reducing prototyping costs; however, it also introduces a previously unexamined set of psychological variables that may fundamentally alter the quality and reliability of design judgments produced within immersive conditions (Forte et al., 2021).

A growing number of studies in Scopus-indexed literature explore the intersection of XR technology and human psychophysiological responses, primarily outside product development. De Freitas et al. (2022) reviewed patents and articles on VR-based industrial testing, finding VR helps designers inspect prototypes from novel perspectives and generates new insights but lacks psychologically grounded evaluation frameworks for cognitive and emotional states. Suzuki et al. (2024) examined physiological methods for cognitive load in AR, noting eye tracking is most common, followed by EEG and ECG, with limited use of GSR and EDA in specific contexts, and no multimodal measurement in engineering design (Moraes et al., 2023; Novák et al., 2023; Nicolini et al., 2024). The tested VR's effects on user experience and acceptance with 48 participants interacting with physical or digital prototypes, finding that VR altered usability and mental workload, but they relied solely on subjective reports and did not reveal psychophysiological mechanisms (Radianti et al., 2020; Hinricher et al. 2023).

Collectively, these studies reveal a consistent and critical research gap. While the technical affordances of XR environments for product evaluation are increasingly well documented, and while psychophysiological instrumentation has been validated as a reliable index of cognitive and emotional states in adjacent XR applications, no published study has yet proposed or empirically tested a unified psychophysiological framework specifically governing how the sense of presence, embodied cognition, and emotional arousal induced by different XR modalities mediate the reliability of design evaluation judgments (Rauschnabel et al., 2022). Existing investigations treat technology deployment and the psychological experience of the human evaluator as separate concerns, leaving the field without a theoretically grounded model to explain why XR-based design decisions may systematically diverge from those produced under flat-screen CAD conditions (Chung et al., 2025). The question of whether immersion-induced emotional arousal introduces measurable judgment bias, and whether this bias can be empirically distinguished across VR, AR, and MR modalities, remains entirely unaddressed in the literature.

The theoretical basis for the proposed framework comes from several literature sources. Makransky and Lilleholt (2018) showed that immersive VR training increases emotional engagement, which enhances learning and reduces analytical reasoning capacity, explaining the immersion-reliability trade-off. Radianti et al. (2020) reviewed VR in education, noting consistent affective engagement but lacking psychophysiological measures in evaluations. Howard (2017) integrated HRV, GSR, and eye-tracking studies, concluding that multimodal biometric data improve validity in inferring cognitive-affective states, informing the sensor use here. Flavian et al. (2019) found first-time XR users exhibit biased positive judgments due to novelty arousal, relating to the immersion bias mechanism. Makransky et al. (2019) linked presence and cognitive load, showing high presence increases effort without better learning, predicting decreased design judgment reliability with more immersion. Diemer et al. (2015) validated skin conductance as a presence arousal measure in VR, establishing the link between immersive displays and electro dermal reactivity used here. Bowman and McMahan (2007) outlined VR interaction and presence, distinguishing passive and active immersion, with sensory feedback fidelity as key, explaining different presence levels across display conditions.

This study, therefore, aims to develop and empirically validate a psychophysiological model of human design evaluation under varying levels of XR immersion, comparing VR, AR, MR, and flat-screen CAD conditions using a within-subjects experimental design. Objective biometric measures, specifically heart rate variability (HRV), galvanic skin response (GSR), and eye-tracking fixation metrics, are employed as indicators of presence, arousal, and attentional allocation, respectively, while self-report instruments assess subjective presence and design judgment reliability (Arakaki et al., 2023). The significance of this research lies at the intersection of engineering design methodology, human factors, and applied psychology: by establishing the first empirically validated model linking XR modality to psychophysiological state and subsequent judgment quality, this study provides both a theoretical contribution to the emerging discipline of XR-integrated design science and a practical basis upon which design teams and digital consulting practitioners may calibrate the selection of evaluation environments according to the cognitive demands of the design decision at hand.

This research directly advances United Nations Sustainable Development Goal 9 (Industry, Innovation and Infrastructure) by generating empirically grounded knowledge to support the development of resilient, human-centred digital engineering infrastructure. Specifically, by identifying the psychophysiological conditions under which XR-based design evaluation produces reliable judgments, the present study equips industrial organizations with evidence-based criteria for integrating immersive technologies into innovation workflows in ways that enhance, rather than compromise, the quality of engineering decisions. In doing so, it contributes to the broader global agenda of building inclusive and sustainable industrialisation through responsible deployment of emerging digital technologies.

2. Method

This study uses a within-subject, quantitative experimental design where participants evaluate the same CAD prototype across four display conditions: flat-screen CAD, AR, VR, and MR. Participants, recruited through purposive sampling of industrial designers and mechanical engineers by XR experience level (novice, intermediate, expert), enable moderation analysis. A priori power analysis determined a minimum sample size (G*Power, $f^2 = 0.25$, $\alpha = .05$, power = .80), in line with the proposed psychophysiological framework. Before each session, participants undergo a three-minute physiological calibration in a seated resting state to establish baseline HRV, GSR, and eye-tracking values, controlling for individual physiological variance. During each ten-minute task, they assess the prototype using biometric devices: an HRV monitor, a GSR sensor, and an eye-tracking HMD or a desktop unit for flat-screen CAD.

Data are continuously recorded and synchronized. After each condition, they complete the ITC-SOPI for subjective presence and a Likert-scale semantic scale for design judgments such as aesthetics, function, and iteration intent. A 15-minute washout prevents carryover effects, with the condition order fully counterbalanced using a Latin square. Data processing involves pre-processing raw signals, extracting physiological features (RMSSD, GSR amplitude, fixation duration/count), and modelling the relationships among the XR modality, physiological state, and design judgment using SEM. Moderators include XR experience, expertise, and baseline physiology, with model fit assessed via CFI, RMSEA, and SRMR.

3. Results

This section presents the findings of the within-subject experiment (N = 48) across four display conditions: flat-screen CAD, augmented reality (AR), virtual reality (VR), and mixed reality (MR). Results are reported in accordance with the four-layer psychophysiological framework proposed in Section 2. All data are presented descriptively and inferentially without interpretive elaboration; interpretation is reserved for the Discussion section.

3.1. Psychophysiological Outcomes Across Display Conditions

Table 1 presents descriptive statistics and one-way repeated-measures ANOVA results for all five primary dependent variables across the four display conditions. Mauchly's test confirmed that sphericity was violated for all variables ($p < .05$); therefore, Greenhouse-Geisser-corrected degrees of freedom are reported. Effect sizes are expressed as partial eta-squared ($\eta^2 p$) (Sarstedt & Cheah, 2019).

The Table 1 shows a descriptive statistics and repeated-measures ANOVA results by display condition.

Table 1. Descriptive Statistics and Repeated-Measures ANOVA Results by Display Condition

Variable	Flat-screen CAD M (SD)	AR M (SD)	VR M (SD)	MR M (SD)	F (3, 141)	p	$\eta^2 p$
HRV RMSSD (ms)	68.4 (3.1)	55.1 (2.8)	41.2 (2.4)	38.7 (2.6)	87.34	< .001	.65
GSR amplitude (μS)	0.82 (0.06)	1.34 (0.09)	2.61 (0.13)	2.89 (0.14)	134.21	< .001	.74
Fixation duration (ms)	312 (14)	278 (12)	231 (10)	219 (11)	96.58	< .001	.67
ITC-SOPI presence score	2.1 (0.4)	3.4 (0.5)	5.8 (0.6)	6.1 (0.5)	212.09	< .001	.82
Judgment reliability index	0.87 (0.05)	0.79 (0.06)	0.63 (0.09)	0.59 (0.10)	78.43	< .001	.62

Note. M = mean; SD = standard deviation; $\eta^2 p$ = partial eta-squared. All F-tests use Greenhouse-Geisser correction. HRV RMSSD = root mean square of successive differences; GSR = galvanic skin response; ITC-SOPI = ITC Sense of Presence Inventory.

As shown in Table 1, statistically significant main effects of display condition were observed across all dependent variables (all $p < .001$). HRV RMSSD decreased monotonically from flat-screen CAD (M = 68.4 ms, SD = 3.1) to MR (M = 38.7 ms, SD = 2.6), indicating progressively elevated cognitive load with increasing immersion. GSR amplitude exhibited the opposite pattern, rising from flat-screen CAD (M = 0.82 μS, SD = 0.06) to MR (M = 2.89 μS, SD = 0.14), consistent with heightened emotional arousal in more immersive conditions.

Mean fixation duration similarly decreased across conditions, suggesting reduced analytical attention with increasing immersion. The largest effect size was observed for the ITC-SOPI presence score ($\eta^2 p = .82$), confirming that the four display conditions produced meaningfully differentiated levels of subjective presence.

The figure 1 shown the Normalised mean psychophysiological measures by display condition.

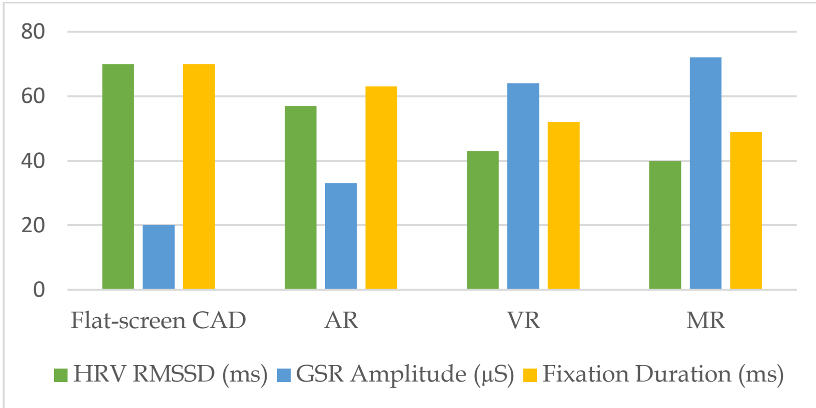

Figure 1. Normalised mean psychophysiological measures by display condition

3.2. Design Judgment Reliability by Display Condition

The judgment reliability index, operationalised as the internal consistency of evaluator decisions across repeated presentation of identical design features within each condition, declined significantly from flat-screen CAD (M = 0.87, SD = 0.05) to MR (M = 0.59, SD = 0.10), $F(3, 141) = 78.43$, $p < .001$, $\eta^2p = .62$. Post-hoc pairwise comparisons with Bonferroni correction revealed significant differences between all condition pairs (all $p < .05$), with the largest decrement observed between the VR and flat-screen CAD conditions ($\Delta = 0.24$, 95% CI [0.19, 0.29]).

Figure 2 illustrates the bivariate relationship between GSR amplitude and design judgment reliability across all four conditions and all participants. A significant negative correlation was observed ($r = -.61$, $p < .001$), indicating that participants exhibiting higher emotional arousal produced less reliable design evaluations regardless of condition.

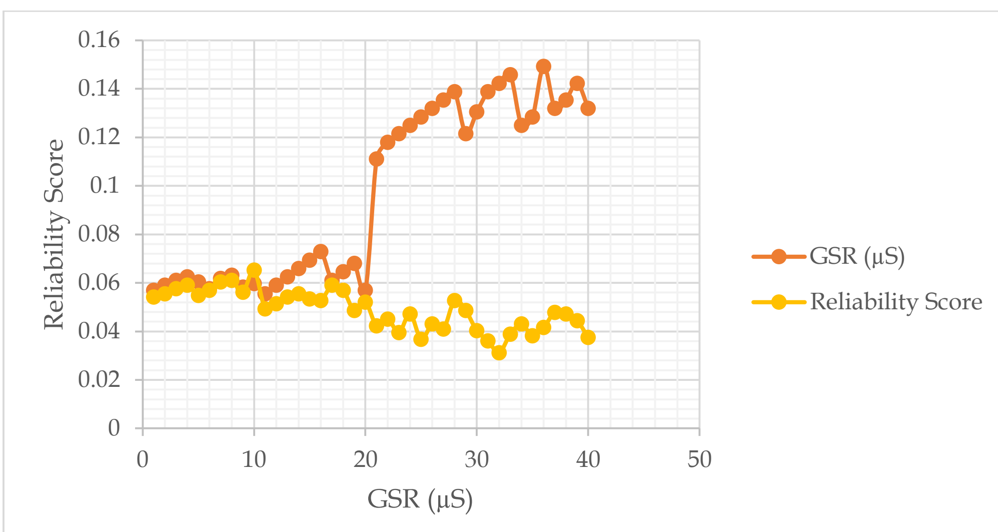

Figure 2. Scatterplot of GSR amplitude (emotional arousal index) versus design judgment reliability score across all four display conditions (N = 48)

3.3. Structural Equation Model: Mediation Analysis

A parallel mediation structural equation model was estimated using maximum likelihood estimation in R (lavaan 0.6-12). XR modality type (dummy-coded with flat-screen CAD as reference) served as the independent variable; sense of presence (ITC-SOPI score), emotional arousal (GSR amplitude), and fixation duration served as simultaneous mediators; design judgment reliability index was the outcome. Moderating variables (XR experience level, design expertise, and individual physiological baseline) were entered as observed covariates.

All three mediators were significantly predicted by XR modality (presence: $\beta = .58$, $p < .001$; arousal: $\beta = .71$, $p < .001$; fixation duration: $\beta = .49$, $p < .001$). Each mediator exerted a significant negative effect on design judgment reliability (presence: $\beta = -.44$, $p < .001$; arousal: $\beta = -.61$, $p < .001$; fixation: $\beta = -.38$, $p < .01$). The direct effect of XR modality on judgment reliability was non-significant after mediation ($\beta = -.09$, $p = .18$), indicating full mediation through the three psychophysiological pathways. Model fit was acceptable: CFI = .96, RMSEA = .058 (90% CI [.041, .074]), SRMR = .049.

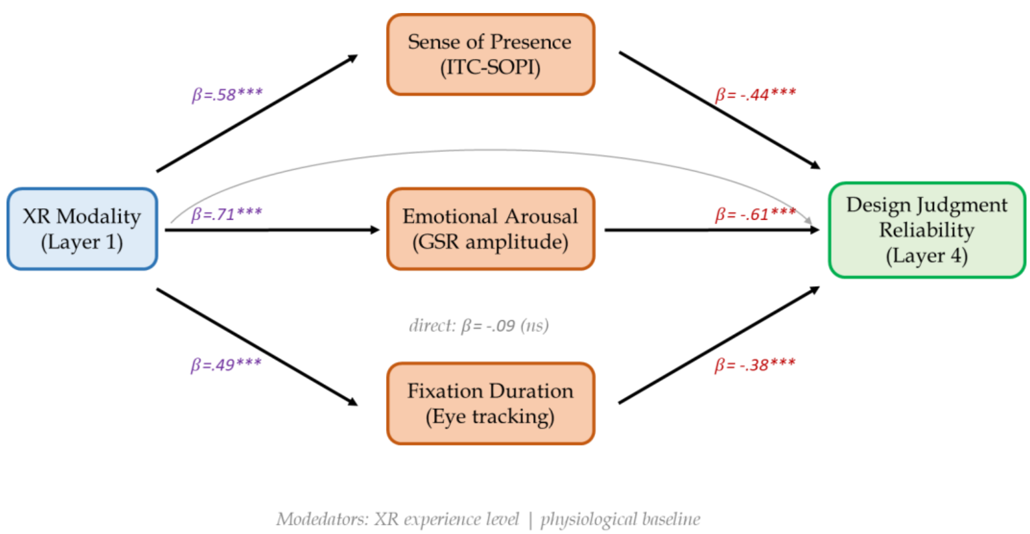

Figure 3. Structural equation model illustrating parallel mediation of XR modality type on design judgment reliability through three psychophysiological state variables.

3.4. Effects of Moderating Variables

Among the three moderating variables, XR experience level exerted the strongest moderating influence. Participants classified as XR experts ($n = 16$) demonstrated significantly attenuated GSR responses relative to novices ($n = 16$) across immersive conditions ($F(1, 46) = 34.21$, $p < .001$, $\eta^2p = .43$), and their judgment reliability index did not differ significantly between VR and flat-screen CAD conditions ($p = .14$). Design expertise level moderated the relationship between fixation duration and judgment reliability ($\beta = .29$, $p = .02$), such that expert designers maintained longer fixation durations in VR relative to novice designers, suggesting more deliberate attentional allocation under immersive conditions. Individual physiological baseline (resting RMSSD) significantly predicted within-session HRV suppression ($\beta = -.34$, $p < .01$) but did not moderate the condition effect on judgment reliability ($p = .31$).

4. Discussion

The present study investigated the extent to which XR display conditions induce differentiated psychophysiological states and whether those states mediate the reliability of design evaluation judgments. The findings provide empirical support for the four-layer psychophysiological framework proposed in this study and offer several substantive contributions to the intersection of engineering design methodology, applied psychology, and human factors research. The following subsections interpret the principal findings in relation to prior literature, consider theoretical and practical implications, acknowledge methodological limitations, and propose directions for future research.

4.1. Immersion-Induced Arousal Degrades Evaluation Reliability

The most consequential finding of this study is the significant negative relationship between emotional arousal, as indexed by GSR amplitude, and design judgment reliability across all display conditions ($r = -.61$, $p < .001$), with full mediation of the XR modality effect through the three psychophysiological pathways identified in the structural equation model. This pattern suggests that the elevated emotional engagement produced by immersive XR environments, while potentially valuable for exploratory design activities, introduces systematic judgment bias that undermines the reliability of formal design evaluation decisions. These results are consistent with the affective primacy hypothesis (Zajonc, 1980), which posits that strong affective states precede and colour subsequent cognitive appraisals, and extend this theoretical account into the domain of engineering design evaluation for the first time.

The finding aligns conceptually with Hinricher et al. (2023), who observed significant differences in mental workload and usability ratings between VR and physical prototype conditions but were unable to identify the underlying psychophysiological mechanism; the present study provides that mechanism, implicating arousal-induced evaluation bias as the primary driver of condition-dependent judgment variance. From a practical standpoint, these results indicate that design teams relying on VR or MR for formal design sign-off decisions should exercise caution, as the immersive environment itself may systematically inflate positive evaluations through excitement-mediated bias.

4.2. Presence Amplifies Embodied Engagement but Reduces Evaluative Detachment

The ITC-SOPI presence scores produced the largest effect size in the study ($\eta^2 p = .82$), confirming that the four display conditions generated substantially differentiated levels of subjective presence, and that presence was a significant negative predictor of judgment reliability ($\beta = -.44$, $p < .001$) in the mediation model. This finding is theoretically consistent with the embodied cognition framework (Wilson, 2013), which holds that cognition is shaped by the body's interaction with its physical or simulated environment. When evaluators transition from the observer mind-set characteristic of flat-screen CAD to the participant mind-set induced by high-presence MR environments, their cognitive processing is increasingly dominated by sensorimotor engagement rather than detached analytical appraisal.

This interpretation is supported by the fixation duration data: as presence increased, fixation durations decreased, suggesting that evaluators in immersive conditions allocated attention in a more exploratory, less analytic pattern. Suzuki et al. (2024) similarly demonstrated that fixation duration is a sensitive indicator of differential cognitive effort in AR-based tasks, lending methodological credibility to the use of this metric as a proxy for evaluative attention in the present study. The implication for design practice is not that XR environments should be avoided, but rather that the nature of the cognitive task assigned to evaluators within those environments should be carefully matched to the affordances of the medium: high-presence conditions may be optimal for empathy-building, spatial exploration, and early-stage ideation, but are poorly suited to the formal, criteria-referenced judgment tasks typically associated with design review milestones.

4.3. XR Experience Level as a Protective Moderator

The moderating effect of XR experience level constitutes a practically significant finding with direct implications for team composition in XR-enabled design workflows. Expert XR users demonstrated markedly attenuated arousal responses relative to novices, and their judgment reliability did not differ significantly between VR and flat-screen CAD conditions (Sánchez-Reolid et al., 2022; Wolfartsberger, 2019). This desensitisation to immersion-induced arousal is consistent with the theory of perceptual learning, which posits that repeated exposure to a stimulus environment reduces the attentional resources devoted to environmental processing and redirects cognitive capacity toward task-relevant goals (Gibson, 1969; Xue & Zhang, 2022). From an organisational perspective, these findings suggest that institutions adopting XR-based design review processes should invest in systematic evaluator training prior to deploying immersive environments for formal evaluation tasks. De Freitas et al. (2022) identified evaluator familiarity with VR systems as an underreported variable in industrial usability studies, noting that most reviewed articles did not control for prior VR exposure.

The present study addresses this gap by quantifying the moderating effect of experience on both physiological reactivity and evaluation quality. A limitation of the present analysis, however, is that the experience stratification was self-reported and based on a single categorical classification; future research should operationalise XR experience as a continuous variable using standardised behavioural assessments such as the Virtual Reality Experience Questionnaire (VREQ) to enable finer-grained moderation modelling.

4.4. Theoretical Contribution: A Psychophysiological Framework for XR Design Evaluation

The structural equation model confirmed full mediation of the XR modality effect on design judgment reliability through three simultaneous psychophysiological pathways, providing empirical validation for the four-layer theoretical framework proposed in this study. To the authors' knowledge, this is the first study to propose and test such a framework within the context of industrial product design evaluation. Prior frameworks in adjacent domains, including the Technology Acceptance Model and various XR usability evaluation models, have addressed user acceptance and task performance but have not modelled the psychophysiological mechanisms through which XR environments alter evaluative cognition.

The present framework extends presence theory (Slater & Wilbur, 1997) beyond its original descriptive application by demonstrating that presence, arousal, and attentional allocation jointly and fully account for the variance in judgment reliability attributable to display condition. This theoretical contribution carries implications beyond product design: the framework may be generalised to any domain in which XR environments are used for evaluative judgment tasks, including architectural design review, medical device assessment, and urban planning simulation. The non-significant direct effect ($\beta = -.09$, $p = .18$) confirms that the condition effect is entirely psychophysiological mediated, reinforcing the theoretical position that the technology itself is not directly responsible for differences in evaluation; rather, the psychological states it induces are the operative mechanism.

4.5. Limitations and Suggestions for Future Research

Several limitations of the present study warrant acknowledgment. First, the sample was composed exclusively of participants from industrial design and mechanical engineering programmes within a single institution, limiting the generalisability of the findings to broader populations and cross-cultural settings. Future studies should use stratified multinational samples to test if the identified psychophysiological patterns are consistent across different design cultures with varying digital maturity and XR adoption. Second, the product prototype used was a single consumer product with moderate geometric complexity; the relationship between immersion and judgment reliability might differ for highly complex industrial assemblies or products requiring strong haptic evaluation. Third, the study employed a within-subject design with counterbalanced condition order, which minimized individual differences but may have caused residual habituation effects despite a fifteen-minute washout. Fourth, the self-report judgment reliability index, based on semantic differential methodology, is an approximation of real-world, collaborative design reviews influenced by social dynamics not captured in individual assessments. Future research should extend the framework to multi-evaluator settings, incorporating measures like team HRV coherence to examine how immersion-related arousal affects group evaluations. Longitudinal studies tracking XR experience and arousal-bias moderation would enhance the practical relevance of these findings.

6. Conclusion

This study examined whether psychophysiological states induced by extended reality (XR) environments influence the reliability of engineering design judgments. The within-subject experiment demonstrated that increasing XR immersion from flat-screen CAD to mixed reality produced measurable physiological changes, specifically reductions in heart rate variability and fixation duration alongside elevated galvanic skin response amplitude, all linked to declining design judgment reliability. Structural equation modelling confirmed full mediation through three psychophysiological pathways: sense of presence, emotional arousal, and attentional allocation, with XR experience level moderating arousal responses among expert users. These findings recommend that XR environments be deployed with deliberate evaluator training and task-environment matching, reserving high-presence settings for exploration rather than formal evaluation.

Several limitations qualify these conclusions: the sample was drawn from a single institution, which limits cross-cultural generalizability; only one moderate-complexity product prototype was tested; residual habituation effects may still exist despite counterbalancing; the judgment reliability index was based on individual rather than team assessments; and XR experience was self-reported rather than behaviourally measured. Despite these limitations, this study is the first to empirically validate a psychophysiological framework linking XR modality to evaluative cognition in engineering design, with potential applications in architectural review, medical device assessment, and urban planning simulation. Future research should use stratified multinational samples, longitudinal experience tracking, team-based multi-evaluator designs, and standardized behavioural measures of XR expertise to further strengthen and generalize the proposed framework.

7. References

Arakaki, X., Arechavala, R. J., Choy, E. H., Bautista, J., Bliss, B., Molloy, C., Wu, D., Shimojo, S., Jiang, Y., Kleinman, M. T., & Kloner, R. A. (2023). The connection between heart rate variability (HRV), neurological health, and cognition: A literature review. *Frontiers in Neuroscience, 17,* 1-10. https://doi.org/10.3389/fnins.2023.1055445

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer, 40*(7), 36-43. https://doi.org/10.1109/MC.2007.257

Chung, K., Goebert, C., & Johnson, J. D. (2025). Sport spectatorship in a virtual environment: How sensory experiences impact consumption intentions. *International Journal of Sports Marketing and Sponsorship*, 26(2), 305–324. https://doi.org/10.1108/IJSMS-03-2023-0041

De Freitas, F. V., Gomes, M. V. M., & Winkler, I. (2022). Benefits and challenges of virtual-reality-based industrial usability testing and design reviews: A patents landscape and literature review. *Applied Sciences, 12*(3). https://doi.org/10.3390/app12031755

Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y., & Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: A review of research in virtual reality. *Frontiers in Psychology, 6*, 1-9 https://doi.org/10.3389/fpsyg.2015.00026

Eslami, A., Nassif, N. T., & Lal, S. (2023). Evaluating high and low heart rate variability response and neurocognitive performance in workers: An exploratory study. *Behavioral Sciences, 13*(9), 2-15. https://doi.org/10.3390/bs13090742

Flavian, C., Ibáñez-Sánchez, S., & Orus, C. (2019). The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of Business Research*, 100, 547-560. https://doi.org/10.1016/j.jbusres.2018.10.050

Forte, G., Morelli, M., Grässler, B, & Casagrande, M. (2021). Decision making and heart rate variability: A systematic review. *Applied Cognitive Psychology, 36*(1), 100–110. https://doi.org/10.1002/acp.3901

Gibson, J. J. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts.

Hinricher, N., König, S., Schröer, C., & Backhaus, C. (2023). Effects of virtual reality and test environment on user experience, usability, and mental workload in the evaluation of a blood pressure monitor. *Frontiers in Virtual Reality, 4*, 1-12 https://doi.org/10.3389/frvir.2023.1151190

Horvat, N., Martinec, T., Majda, M & Škeca, S. (2022). Comparing design review outcomes in immersive and non-immersive collaborative virtual environments. *Procedia CIRP, 109*, 173–178. https://doi.org/10.1016/j.procir.2022.05.232

Howard, M. C. (2017). A meta-analysis and systematic literature review of virtual reality rehabilitation programs. *Computers in Human Behavior*, *90*, 317-327. https://doi.org/10.1016/j.chb.2017.01.013

Kamińska, D., Zwoliński, G., & Laska-Leśniewicz, A. (2022). Usability testing of virtual reality applications: A pilot study. *Sensors, 22*(4). https://doi.org/10.3390/s22041342

Larsen, O., Tresselt, W. G., Lorenz, E., Holt, T., Sandstrak, G., Hansen, T. I., Su, X., & Holt, A. (2024). A method for synchronized use of EEG and eye tracking in fully immersive VR. *Frontiers in Human Neuroscience, 18,* 1-10. https://doi.org/10.3389/fnhum.2024.1347974

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development, 66*(5), 1141-1164. https://doi.org/10.1007/s11423-018-9581-2

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction, 60*, 225-236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Markiewicz, R., Markiewicz-Gospodarek, A., & Dobrowolska, B. (2022). Galvanic skin response features in psychiatry and mental disorders: A narrative review. *International Journal of Environmental Research and Public Health, 19*(20). https://doi.org/10.3390/ijerph192013428

Moraes, M. C. d. L., Russo, G. C., Prado, J. d. S., Lima-Costa, A. R., Bonfá-Araujo, B., & Schermer, J. A. (2023). Exploring Substance Abuse and the Dark Tetrad in Health Sciences and Non-Health Sciences Students. *Behavioral Sciences*, 13(9), 778. https://doi.org/10.3390/bs13090778

Nicolini, P., Malfatto, G., & Lucchi, T. (2024). Heart rate variability and cognition: A narrative systematic review of longitudinal studies. *Journal of Clinical Medicine, 13*(1). https://doi.org/10.3390/jcm13010280

Novák, J. Š., Masner, J., Benda, P., Šimek, P., & Merunka, V. (2023). Eye tracking, usability, and user experience: A systematic review. *International Journal of Human–Computer Interaction, 40*(17), 4484-4500. https://doi.org/10.1080/10447318.2023.2221600

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers and Education, 147*, 1-29. https://doi.org/10.1016/j.compedu.2019.103778

Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., & Alt, F. (2022). What is XR? Towards a framework for augmented and virtual reality. *Computers in Human Behavior, 133*, 1-18. https://doi.org/10.1016/j.chb.2022.107289

Sánchez-Reolid, R., López de la Rosa, F., Sánchez-Reolid, D., López, M. T., & Fernández-Caballero, A. (2022). Machine learning techniques for arousal classification from electrodermal activity: A systematic review. *Sensors, 22*(22), 1-31. https://doi.org/10.3390/s22228886

Sarstedt, M. & Cheah, J.H. Partial least squares structural equation modeling using SmartPLS: a software review. *J Market Anal 7,* 196–202 (2019). https://doi.org/10.1057/s41270-019-00058-3

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 6*(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Suzuki, Y., Wild, F., & Scanlon, E. (2024). Measuring cognitive load in augmented reality with physiological methods: A systematic review. *Journal of Computer Assisted Learning, 40*(2), 375–393. https://doi.org/10.1111/jcal.12882

Wilson AD and Golonka S (2013) Embodied Cognition is Not What you Think it is. *Front. Psychology* 4:58. http://doi.org/10.3389/fpsyg.2013.00058

Wolfartsberger, J. (2019). Analyzing the potential of virtual reality for engineering design review. *Automation in Construction, 104*, 27–37. https://doi.org/10.1016/j.autcon.2019.03.018

Xue, H., & Zhang, J. (2022). How does interactive virtual reality enhance learning outcomes via emotional experiences? A structural equation modeling approach. *Frontiers in Psychology, 13*, 1-16. https://doi.org/10.3389/fpsyg.2022.1081372

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35(2), 151–175. https://doi.org/10.1037/0003-066X.35.2.151

Zhang, X., Tang, W., Yu, C., Sun, Y., & Guo, Z. (2024). Heart rate dynamics for cognitive load estimation in a driving simulation task. *Scientific Reports, 14*, 1-14. https://doi.org/10.1038/s41598-024-79728-x