Design of a Personalised AI Coaching Assistant for Occupational Health and Safety

Jonathan Vitale[1(✉)], Shlomo Berkovsky[1], Shun Takeuchi[1,3], Amin Beheshti[2], Kexuan Xin[1], Junya Saito[3], and Sosuke Yamao[3]

[1] Australian Institute of Health Innovation, Macquarie University, Macquarie Park, Australia
{jonathan.vitale,shlomo.berkovsky,kexuan.xin}@mq.edu.au
[2] Centre for Applied Artificial Intelligence, Macquarie University, Sydney, Australia
amin.beheshti@mq.edu.au
[3] Artificial Intelligence Laboratory, Fujitsu Research, Kanagawa, Japan
{s.takeuchi,saito.junya,yamao.sosuke}@fujitsu.com

Abstract. This work presents a personalised AI coaching system to enhance occupational health and safety using a bottom-up design approach inspired by Lean UX and Agile principles. Leveraging Large Language Models and Computer Vision, the pilot integrated automated reporting and role-playing simulations to address safety challenges. Prototyping with existing and adapted AI tools demonstrated feasibility, with positive feedback from managers highlighting its potential to improve compliance and the need for staff involvement. The study underscores AI's role as a collaborative coaching mediator and the effectiveness of bottom-up design in aligning solutions with user needs, while naturally integrating persuasive technology principles.

Keywords: Human-Centred Design · Personalised Coaching · Persuasive Technologies · Collaborative AI · Occupational Health and Safety Agile Methodologies · Role-Playing · LLMs

1 Introduction

Occupational Health and Safety (OHS) safeguards workers' physical and psychological well-being [39]. Warehouses pose heightened risks due to hazardous materials and regular use of heavy machinery, e.g. forklifts [23], that could potentially lead to severe injuries or fatalities [37]. Other common incidents include slips, trips, and improper manual handling [32], which may cause musculoskeletal disorders, chronic conditions, and operational disruptions [5,34]. A proactive safety culture with continuous training and strict adherence to protocols reduces risks [23,37]. While all employees must follow OHS guidelines, OHS managers

K. T. Win et al. (Eds.): PERSUASIVE 2025, LNCS 15711, pp. 3–17, 2025.
https://doi.org/10.1007/978-3-031-94959-3_1

carry the primary duty of identifying, mitigating, and eliminating safety hazards [30]. Despite this, hazards can still be overlooked, rules ignored for convenience, and human errors made [39].

AI can be used to mitigate workplace risks by processing multimodal data to detect hazards and unsafe behaviour [1,24,31,39,47]. Automating tasks like document review and surveillance reduces managerial workload, while Large Language Models (LLMs) enhance reporting [33], cause-effect analysis [4], and natural interactions. As Socially Intelligent Agents (SIAs) [12,13], LLMs can understand and respond to social behaviour [27,48]. Designed as socio-technical tools [46], they integrate with ubiquitous technologies for situational awareness, recognising specific users, preferences, mental states, actions, and needs [49]. This capability not only fosters seamless human-machine collaboration but also supports the delivery of personalised assistance, defined here as tailoring experiences to the unique preferences, behaviours, and needs of individual users [8].

These capabilities extend to the identification of workplace issues impacting OHS, reporting them efficiently, and assisting with the implementation of personalised solutions to mitigate risks before incidents occur [35]. Additionally, these technologies can be used to *persuade* users toward safer and more compliant behaviours. For instance, AI can deliver timely and context-sensitive recommendations to encourage adherence to OHS protocols via the *principle of suggestion* [16], gradually fostering a culture of compliance. However, poorly designed AI tools can introduce unintended risks or exacerbate existing issues, potentially increasing the volume of OHS incident reports [6]. This underscores the critical importance of adhering to proper design processes that ensure AI systems are effective and align with desired outcomes, a goal that is central to the aim of this work.

This work lies at the intersection of OHS and persuasive technologies (PT) design, leveraging recent advancements in LLMs. To the best of our knowledge, this is the first study to integrate LLMs and PT in addressing OHS challenges. PT design methodologies typically adopt a top-down approach [2,11,15], relying on predefined behavioural change interventions, problems or selected persuasive principles, even when stakeholders are included in participatory activities [22,41] or bottom-up elements are integrated in the design process [26]. In contrast, we take a different perspective. We employ a *full* bottom-up design methodology inspired by Lean UX and Agile methodologies, which starts with uncovering user needs to enhance their experience in the considered context, allowing insights to emerge organically and inform the design process without being constrained by a predefined behavioural intervention or problem. This ensures that solutions remain grounded in the complexities of real-world contexts, addressing practical challenges faced by users.

We present the design process followed and its initial outcomes, demonstrating how a bottom-up approach, previously applied in social robotics [42], can guide the development of user-centred solutions aligned with PT design principles. We specifically focus on design of PT incorporating coaching by human experts, such as OHS managers, and position AI as coaching mediator to augment human expertise, fostering collaboration and enhancing OHS compliance. Hence, the work contributes PT research in several ways. *First*, we explore the adaptation of a bottom-up design to PT, detailing its phases and activities used to generate outcomes, while demonstrating its effectiveness in uncovering user needs and naturally integrating persuasive principles. *Second*, we present an early-stage pilot of a personalised AI coaching assistant for the OHS domain, showcasing how the capabilities of LLMs can enable persuasive design principles, such as simulation and rehearsal, to enhance coaching effectiveness. *Third*, we propose a new interpretation of PT for coaching applications, positioning AI as a mediator that supports and augments human expertise, enabling collaborative and context-aware behavioural change interventions.

2 Design Process

This work originates from a collaborative project with Fujitsu Japan, aimed at developing AI-based personalised coaching solutions in an industrial setting. Such solutions leverage AI technologies, including Computer Vision, for activity recognition and aim to integrate LLMs for natural, socially intelligent, and tailored coaching. To design them, we adapted Tonkin et al.'s methodology [42] originally conceived for designing innovative commercial social robotics applications with positive user experience. This methodology was chosen for its adaptability and flexibility in refining design problems iteratively.

The methodology's *adaptability* lies in its integration of technology, user context, and situated interaction, allowing it to adjust flexibly to different use cases and technologies by reconfiguring these key components as needed. Originally conceived for social robots, its principles extends to SIAs embedded in physical settings through cameras and smart devices. This integration has proven beneficial for PT design [28], yet few studies detail how to achieve it through specific design activities [26], highlighting the need for more works like the one reported here. In addition, the methodology's *flexibility* allows it to evolve within the same project as stakeholder insights emerge, fostering exploration to identify relevant use cases based on observed user needs. Combining Lean UX [20] for creating minimum viable products (MVPs) with Agile Science [21] for testing and refinement ensures that the technology is both practical and evidence-based.

The following paragraphs outline the design process and its outcomes[1], limited to the first five phases of the methodology [42] and culminating in a preliminary pilot assessed for technical feasibility and user feedback.

Phase 1: Define the Challenge. The first phase focussed on identifying a suitable context for implementing the desired personalised AI coaching. Warehouse management emerged as the promising context for developing the AI coaching solution. To guide the subsequent design, this phase ended with a situated

[1] Design deliverables, pilot, and prompts used for fast prototyping LLM's features can be found in the online addendum, `https://osf.io/kd2pq/?view\_only=782b7dde4aef41228eec3f6b1dec0121`.

How-Might-We (HMW) statement, a concise question that frames challenges in a solution-oriented manner [3,40]. The resulting HMW statement was:

> *"How might we use AI to enhance the work experience of managers and staff at Fujitsu's warehouse by delivering personalised coaching that supports their individual needs?"*

This HMW statement points the design to the targeted AI technology, specifying the context and users, while keeping the desired outcome, i.e., enhancing the work experience, broad enough to encourage exploration of the problem space. At this stage, specific user needs and behavioural interventions deliberately remain unknown, following a bottom-up approach to allow flexibility in exploring how AI can address emerging needs within the chosen context.

Phase 2: Observe. In this phase, sample warehouse surveillance footage was obtained and processed with Computer Vision. The footage was automatically annotated with bounding boxes and labels for people and objects, depicting staff activities, ranging from basic actions like walking and bending to more complex scenarios, e.g., "close to a moving forklift". The footage familiarised the team with the warehouse environment in preparation for meetings with managers.

Subsequently, the team met with the warehouse managers to deepen the understanding of critical operations, focussing on internal logistics, item dispatch, and risk management. While limited needs were identified in logistics and dispatch, significant opportunities emerged in risk management, also aligning well with the motivating use-case for AI coaching. The team then conducted an on-site visit, to obtain detailed insight on warehouse risks and past incidents. By the end of this phase, the team gained an awareness and understanding of the warehouse environment. Key insights around staff training and risk management were documented and reinforced in discussions with OHS managers, forming solid foundations for the subsequent design phases.

Phase 3: Form Insights. The insights and notes gathered during the observation phase were organised in an Empathy Map [36,43], with each detail placed in the appropriate quadrant: what users see, think, feel, and do. This structure helped the team empathise with users' needs, emotions, experiences and actions. Each insight was then classified as a 'gain' (positive aspect to preserve) or 'pain point' (area to improve/mitigate), or was kept neutral. The notes in each quadrant of the Empathy Map were re-organised using affinity mapping [36], which organised similar insights into themes, highlighting key strengths and challenges in warehouse operations.

An analysis of the identified themes followed. On the positive side, warehouse staff benefitted from continuous training and licensing, supported by a new national program certifying machinery competency. Regular use of machinery was found to reduce risks, whereas incident investigations effectively prevented repeated issues. Encouraging staff to acknowledge mistakes promoted accountability and reduced escalation risks. However, several challenges were identified. Personal differences among staff, over-confidence, and complacency with OHS rules contributed to breaches. Annual training sessions and reading materials were found inadequate, with staff often prioritising convenience over safety. On-site monitoring was time-consuming for OHS managers and investigating incidents added burden. The analysis also uncovered several tension points, specific areas of conflict creating opportunities for solutions. While clear OHS guidelines had been available to staff, these were often disregarded due to low perception of risks. Likewise, the confidence gained from mandatory training sometimes led to complacency, undermining safety practices. Finally, the new training program was praised by managers but deemed insufficient due to its low frequency, highlighting the need for ongoing reinforcement.

This phase concluded with the creation of an end-to-end journey map [40], focussing on the activities, emotions, and challenges of the OHS managers and highlighting risk management operation points with a low user experience.

Phase 4: Frame Opportunities. Each identified low user experience point is an opportunity for improvement. Thus, we crafted a corresponding HMW statement to highlight these opportunities for targeted solutions. This led to the following HMW statements: (i) How might we reinforce OHS compliance knowledge among staff to lead to fewer OHS breaches and ensure that OHS managers feel valued and heard?; (ii) How might we coach staff to prioritise OHS-compliant and safe behaviour, discouraging them from choosing convenient but unsafe actions?; (iii) How might we ensure staff remains focussed on OHS-compliant behaviour and avoids overconfidence while performing tasks in the warehouse?

The three HMW statements were consolidated into one *contextualised statement*, incorporating the use of AI as a guiding constraint. This ensured alignment with the project objectives, while informing the subsequent ideation phase:

> *"How might we use AI to assist warehouse managers in coaching staff to prioritise OHS-compliant and safe behaviour over convenience, while reinforcing their OHS knowledge, maintaining focus on safe practices, and fostering attentiveness to their activities, thereby reducing OHS breaches?"*

Phase 5: Brainstorm. The contextualised HMW statement inspired ideas for brainstorming, where the team was encouraged to propose even unconventional ideas to foster creativity. The session emphasised collaboration, with the team members building on each other's ideas to explore diverse perspectives and refine concepts into innovative solutions. Over 70 ideas were generated and analysed, yielding eight overarching themes.

The *"Data Integration and Processing"* theme focussed on ideas leveraging multimodal data to prevent and identify OHS breaches. The theme of *"Data Explainability"* encompassed ideas aimed at providing managers with insights on the causes of the detected issues, enhancing their awareness and ability to resolve underlying problems. Ideas under the *"Automatic Generation of OHS Materials"* theme proposed automating the creation of tailored OHS content and reports, reducing manual effort and aligning outputs with user profiles.

The *"AI as Coaching Mediator"* theme positioned AI as a facilitator between managers and staff, enabling personalised coaching simulations and improving communication. The theme of *"Managers in Control of AI"* highlighted ideas empowering OHS managers to provide ongoing feedback and ensure alignment with their needs. Meanwhile, the *"Real-time Activity Monitoring"* theme covered ideas around using live surveillance video to detect risky behaviour and signs of employee fatigue, enabling timely notifications. Finally, the *"Gamification and OHS Compliance Rewards"* theme explored ideas incorporating rewards to promote adherence to OHS protocols, while the theme of *"Centralised Collaborative Platform"* envisioned a shared network of warehouses to share best practices, coaching recommendations, and risk management resources.

Following the analysis, these themes were ranked based on technical feasibility and their potentials in mitigating OHS incidents. The highest-ranked concepts and ideas were used to guide the design of a pilot addressing the identified needs. This marked the final step of the design process reported in this work, aligned with the incremental approach that emphasises iterative design cycles incorporating evaluation and preliminary feedback. The remaining steps outlined by Tonkin et al. [42] were beyond the scope of this paper and will be reported in future works.

3 HIRO, the AI Coaching Assistant

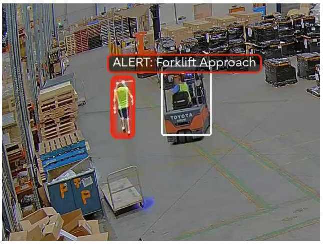
Building on the previously generated ideas, we identified and refined concept solutions addressing the challenges and needs uncovered by the design process. These were materialised into a high-fidelity pilot AI coach called HIRO (Hazards, Incidents, and Risks Operations), designed to assist OHS managers (see Fig. 1). While the pilot resembled the intended coach's functionality and interactions, its features were emulated to demonstrate how the coach would work in practice.

Fig. 1. HIRO.

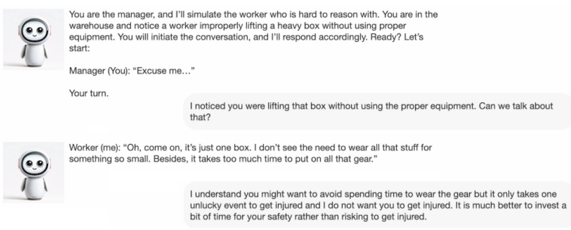
3.1 Pilot Features

The designed pilot incorporated three key features tailored to address critical aspects of warehouse OHS risk management:

Data Integration and Reporting: A preliminary analysis of the current situation is crucial for identifying safety management needs, enabling OHS managers to assess warehouse issues and propose effective solutions [44]. As such, this feature integrates warehouse data selected by managers to identify issues and uncover root causes. Managers can query HIRO about specific problems or needs, and receive insights based on contextual data analysis. Additionally, HIRO includes an automatic report-generation tool that presents visual findings, such as plots or surveillance shots. These tools streamline reporting, enhance *situational awareness*, and provide actionable insights to support decision-making.



(a) Human-Sensing technology.



(b) Role-Play Simulation with LLM.

Fig. 2. Examples of prototyped features leveraging available AI technologies.

Role-Playing Coaching Simulations for Staff: Frequent employee training is essential for raising the safety standards and ensuring employees are educated on reducing risks [44]. Staff needing OHS training can be identified through the data integration feature and assigned to coaching interventions. HIRO generates personalised scenarios following OHS managers' requests for customised coaching simulations addressing specific training needs. Staff interacts with these simulations through natural dialogue and receives tailored improvement recommendations. The manager receives a report too, which includes additional recommendations and suggestions for future staff training.

Role-Playing Simulations Vignettes for Managers: OHS managers can use HIRO to simulate and rehearse meetings with staff. HIRO can simulate worker personas with realistic attitudes and reactions. These scenarios test managers' professional and interpersonal skills, helping them refine strategies for addressing non-compliance. Upon completing the simulation, HIRO evaluates the manager's performance and offers recommendations for improvement.

These features provide *personalised assistance* [8] by tailoring role-playing simulations to the coaching needs of OHS managers, customising scenarios, adapting worker personas, and generating tailored reports to support both managerial and staff development.

Building on the key features outlined in the proposed pilot, we explored existing AI technologies to fast prototype the core functionalities needed for their implementation, following the Agile principle of rapid iterations. Figure 2 exemplifies two technologies used to demonstrate the pilot's functionalities. Figure 2a showcases Fujitsu's human sensing AI, Fujitsu Kozuchi for Vision [17,18], applied to pre-processed warehouse video footage. The system uses bounding boxes to identify people and warehouse machinery, labelling them with corresponding actions. Simple actions, such as standing, bending, or walking, represent basic movements or poses of staff. In contrast, complex actions, such as detecting unsafe proximity between a staff member and a moving forklift, involve interactions between multiple entities and rely on editable inference rules. These rules are managed through a user interface allowing users to define, compose, and modify rules, including adding support for recognising new entities. Once recognised, these actions were further represented as Action-Scene Graphs [25], providing a structured and detailed representation of warehouse activities, which were then utilised by a Retrieval-Augmented Generation (RAG) system to efficiently retrieve and inject relevant information to an LLM for the generation of contextual insights [19]. This representation enhances the system's ability to analyse and interpret multimodal data, enabling the identification of unsafe behaviours and supporting timely reporting to warehouse staff and managers.

Figure 2b shows an excerpt from a GPT-4-generated dialogue where the manager can practice interpersonal skills by interacting with a simulated worker, handling resistance to encourage OHS compliance. The model was used as-is, with task-specific prompt engineering to align with the application domain. Although the initial prompts did not include domain-specific knowledge or worker profiling details, the LLM was shown to be adaptable to craft scenarios testing a manager's communication and interpersonal skills. The prompt included an initial interaction, where the LLM asked questions to personalise the coaching scenario based on the manager's needs. The model then generated a unique tailored vignette and concluded with actionable recommendations for improvement. This proof-of-concept highlighted the potential of LLMs to facilitate coaching across domains, with future implementations aiming at enabling dynamic prompt personalisation and reducing the need for extensive prompt engineering.

3.2 OHS Managers' Feedback

The feedback of OHS managers on the pilot was generally positive, transitioning from a neutral starting mood to enthusiasm, as they realised the potential of Computer Vision and LLMs to their field. One manager remarked that multimodal data integration for incident analysis would simplify the way incidents are investigated at the warehouse. The pilot's role-playing dialogues also resonated with OHS managers, who highlighted its realism and even noted how accurately the simulated responses mirrored workers' reactions to OHS enforcement.

A senior OHS manager provided constructive suggestions, emphasising the importance of involving staff in the design process to ensure acceptance and engagement. He cautioned that introducing the technology without their participation might lead to resistance, as staff could perceive it as a surveillance tool rather than a beneficial resource. Aligned with this perspective, privacy by design principles have been embedded from the outset, ensuring that transparency and user control mitigate privacy concerns while enhancing trust and user experience [45]. To further foster a sense of ownership, he recommended involving team leaders in future design iterations, highlighting that demonstrating the technology's impact on reducing incidents would help staff recognise its value. He noted: "*If we can show staff how this technology reduces incidents and makes the workplace safer, they'll see the value themselves*". Additionally, he suggested using concrete metrics, such as monthly incident reduction, to effectively communicate the benefits of the pilot. Such metrics not only illustrate its impact but also serve as an efficacy measure for the PT, reinforcing its role in reducing non-compliant behaviour. Another key concern was false positives in incident detection, potentially increasing OHS managers' workload. The senior manager noted: "*Incorrect data would mean more work, as we'd have to manually review and dismiss false incidents*". The team clarified that the pilot would generate aggregate statistics to minimise individual errors and include a feedback mechanism for managers to flag inaccuracies, refining AI performance through human-in-the-loop input.

Overall, OHS managers endorsed the proposed pilot and supported staff involvement in future design phases and testing, reflecting a shared vision for the AI coach and its potential to enhance OHS attitudes and behaviours.

4 Retrospective Analysis and Key Learnings

Building on managers' feedback, this design work ensured alignment with practical needs, with an Agile approach enabling iterative improvements. The methodology advances through the *Experiment* and *Measure* phases [42], using user experiments to assess impact via qualitative and quantitative measures. While these phases extend beyond this paper, the strong foundations laid so far ensure future development follows a robust design framework. This section reflects on the reported design activities, using retrospective analysis [38] to extract key learnings that contribute more broadly to the field of PT.

4.1 Bottom-Up Emergence of PT Principles

The work demonstrates the effectiveness of a bottom-up design, contrasting the top-down methods traditionally employed in PT. Top-down approaches use expert-defined interventions and tailored persuasive design principles to drive behavioural change, but may overlook user needs or the causes of the behaviour targeted for change. In contrast, our process did not initiate with a fixed intervention, goal, or principles, focussing instead on exploring the potential of the AI coaching technology for the chosen context. This open-ended approach allows the design to evolve through iterative stakeholder engagement and insights, gradually uncovering pain points and tailoring innovative solutions.

The bottom-up methodology offers notable advantages. Its adaptability enables exploration and selection of use cases aligned with emergent needs, such as improving OHS compliance. The process naturally uncovered the root causes of the targeted behaviour, ensuring it addresses not only the behavioural goal but also systemic issues, such as inefficiencies in monitoring, shortcomings in training methods, or the need for managers to instil OHS compliance more effectively.

The outcomes of the design process notably aligned with established PT principles [9,10,16], which were not imposed top-down but emerged organically. For example, the *reduction* principle can be observed in features enabling OHS managers to streamline data integration and processing to identify safety risks efficiently, and deliver personalised OHS coaching with the AI acting as a mediator. Similarly, *tailoring* is reflected in customised coaching simulations and reports, dynamically adapted to the user's role and needs. The design incorporates the *simulated cause-and-effect* and *simulated environments* principles, offering realistic role-playing vignettes for managers and staff. These allow users to rehearse behaviours, observe decision outcomes in real-time, address knowledge gaps, and improve professional and interpersonal skills. Moreover, the *suggestion* principle enhances decision-making by delivering timely and actionable recommendations grounded in integrated data or following the role-playing interactions. To encourage compliance, the *surveillance* principle uses ubiquitous AI to enable managers to monitor staff OHS behaviour, promoting compliance through awareness of being observed. Finally, the *authority* principle strengthens adherence by framing simulations as directives from managers and referencing OHS rules, reinforcing the importance of training and compliance.

What does this mean for the PT community? In some cases, traditional top-down design may be an effective way to address the target behaviour, specifically when the underlying causes are well-understood and clear strategies for intervention can be identified. As an analogy, take a 'supervised' machine learning approach, where the ground truth is known and the training process is guided by existing labels. Similarly, when use case-specific knowledge is not required and general pre-existing knowledge is sufficient, selecting PT principles and ideating solutions around them can be appropriate. However, when the target behaviour is unclear, the underlying causes are not fully understood, or it is uncertain if a behaviour change is beneficial for the considered context, a bottom-up approach offers a solid complementary solution. Analogous to 'unsupervised' machine learning, where patterns and insights emerge without labels, bottom-up approach facilitates an in-depth exploration of users' needs and pain points without relying on predefined assumptions. For example, in our work, the holistic investigation of the warehouse environment revealed that coaching interventions were needed not only for warehouse staff but also for OHS managers. Managers required solutions to improve their situational awareness and develop interpersonal skills, enabling them to address staff non-compliance more effectively.

By focussing on user-centred design activities, the process shifts from applying pre-selected PT design principles to understanding the users' context and problems. Only after a target behaviour is identified, can PT principles be layered onto the design process to refine solutions effectively. This adaptive approach allows the PT community to explore a broader solution space, grounding interventions in user insights, fostering innovation, and developing technologies that are more attuned to real-world needs and user expectations.

4.2 Collaborative Human-and-AI Coaching

This work brings to the fore a new perspective on the role of coaching PT. Traditionally, PT are designed to function independently of human coaches or experts, directly engaging with users to drive behavioural change [16]. In such cases, the expertise and strategies of human coaches or experts are distilled and embedded into PT to increase scalability and eliminating the need for continuous human input [16]. While this is advantageous in contexts where experts

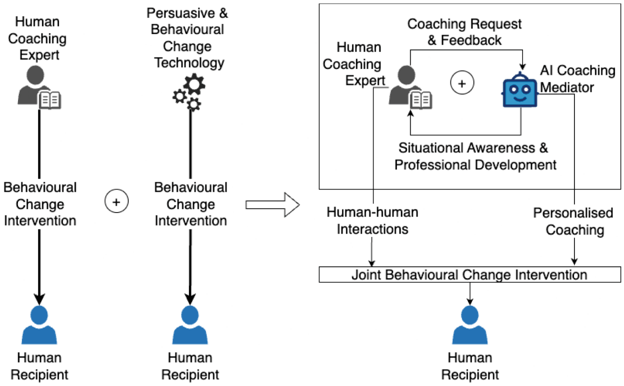


Fig. 3. Combining human-and-AI coaching fosters collaboration and enhances behavioural change, while maintaining workplace interactions.

are unavailable or can be safely removed, it may be a notable limitation, e.g., in the OHS setting explored by this work. OHS managers play a pivotal role in delivering guidance, as well as fostering trust and accountability. By replacing OHS managers, PT may alienate staff and diminish the impact of behavioural interventions, as interpersonal dynamics are critical for ensuring OHS compliance [29]. In general, fully delegating behavioural interventions to PT in high-risk and mission-critical domains may be unacceptable and have legal consequences.

As shown in Fig. 3, our design yielded an alternative solution, positioning AI coaching not as a replacement to humans, but as a collaborative mediator augmenting the humans. This positions AI, beyond the means to generate and deliver personalised coaching, also as a resource that enhances the human coach's knowledge, awareness, and skills, thus, more effectively guiding and influencing the recipient of the coaching interventions through direct interpersonal interactions [29]. In the context of OHS, this dual effort can enhance compliance and ultimately reduce OHS incidents, a key evaluation metric for the PT.

The versatility of this framework may extend to other coaching domains, e.g., education, where AI can deliver personalised learning to students while providing teachers with insights on classroom dynamics and student needs. This can help educators adjust strategies to improve learning outcomes. Our work shows how PT embedding expert coaching can elevate recipient-focused tools into holistic collaborative support technologies. By integrating human expert and AI, we can enhance interventions through a dynamic, feedback-driven partnership.

4.3 Limitations and Future Directions

This work highlighted the value of a bottom-up design methodology in uncovering PT principles organically and positioning AI as a collaborative mediator for AI coaching. Despite establishing a robust foundation, several limitations, which are at the same time opportunities for future research, should be discussed.

First, our participatory design process was limited to the research and development team and OHS managers, without direct involvement from broader warehouse staff. This decision was driven by the need to initially establish a clear understanding of managerial workflows and priorities before expanding the stakeholder engagement. Concerns around potential resistance to new technologies at early development stages also influenced this decision. Future iterations will actively involve broader staff through participatory sessions, to incorporate their unique perspectives and ensure the system addresses their needs. Such a phased approach enables robust foundational development, while creating opportunities to refine the system based on staff feedback and fostering engagement and ownership [26].

Another limitation is the absence of a comprehensive user study evaluating HIRO's effectiveness in tackling OHS non-compliance. This aligns with the current focus on the design process rather than empirical validation. The chosen design methodology [42] includes subsequent experimental phases that build on the design outcomes presented here. These will involve structured user studies with quantitative metrics, such as comparing the accuracy of OHS managers in identifying non-compliance in video footage with or without the AI technology, task performance times, and more. Ultimately, short- and long-term trends in OHS incidents before and after deploying the complete solution will be analysed. Qualitative assessments, including user experience evaluation to understand satisfaction and usability, as well as think-aloud protocols [7,14] to capture real-time thoughts during interactions, will complement the quantitative metrics.

Finally, an important limitation of this paper refers to the lack of detailed technical descriptions. These include the implementation of the computer vision and LLM features, as well as their integration underpinning the proposed solution. This omission was a deliberate choice to ensure the paper focusses on the design methodology and its broader implications for the PT community. The LLM prompts used to generate the sample coaching simulations presented in the pilot are available in the online addendum referred to in Sect. 2, while the technical details of the full implementation will be presented in future publications. The limited presentation of the design outcomes was primarily dictated by page constraints, with these outputs also accessible in the online addendum.

5 Conclusions

This work highlights the effectiveness of a bottom-up design approach in identifying user needs and addressing OHS challenges in warehouses. The AI coach, HIRO, supports OHS managers in fostering compliance and safety culture, with positive feedback validating the participatory approach's role in user trust and adoption.

A retrospective analysis underscored the value of AI coaching as a collaborative mediator that enhances rather than replaces human coaches. AI provides tailored insights, reinforcing interpersonal interactions and strengthening OHS managers' capabilities. Additionally, the bottom-up design naturally integrated

PT principles, enabling adaptive, realistic coaching interventions without predefined behaviours or goals.

Acknowledgments. This work is based on results obtained by Fujitsu Macquarie AI Research Lab.

References

1. Akinsemoyin, A., Awolusi, I., Chakraborty, D., Al-Bayati, A.J., Akanmu, A.: Unmanned aerial systems and deep learning for safety and health activity monitoring on construction sites. Sensors **23**(15), 6690 (2023)
2. Almohanna, A.A.S., Meedya, S., Vlahu-Gjorgievska, E., Win, K.T.: Exploring user experiences with a persuasive mhealth app for breastfeeding: an empirical investigation. International J. Hum.–Comput. Interact. 1–18 (2024)
3. Asano, Y.: Defining the problems solution to lead to the ideation phase: a case study on the use of how might we.... In: International Conference on Human-Computer Interaction, pp. 27–37. Springer, Cham (2023)
4. Ashwani, S., et al.: Cause and effect: can large language models truly understand causality? In: Proceedings of the AAAI Symposium Series, vol. 4, pp. 2–9 (2024)
5. Basahel, A.M.: Investigation of work-related musculoskeletal disorders (MSDs) in warehouse workers in Saudi Arabia. Procedia Manuf. **3**, 4643–4649 (2015)
6. Cebulla, A., Szpak, Z., Knight, G.: Preparing to work with artificial intelligence: assessing WHS when using AI in the workplace. Int. J. Workplace Health Manag. **16**(4), 294–312 (2023)
7. Charters, E.: The use of think-aloud methods in qualitative research an introduction to think-aloud methods. Brock Educ. J. **12**(2) (2003)
8. Chen, J., et al.: When large language models meet personalization: perspectives of challenges and opportunities. World Wide Web **27**(4), 42 (2024)
9. Cialdini, R.B.: The science of persuasion. Sci. Am. **284**(2), 76–81 (2001)
10. Cialdini, R.B.: Influence: Science and Practice, vol. 4. Pearson Education, Boston (2009)
11. Colusso, L., Do, T., Hsieh, G.: Behavior change design sprints. In: Proceedings of the 2018 Designing Interactive Systems Conference, pp. 791–803 (2018)
12. Dautenhahn, K.: The art of designing socially intelligent agents: science, fiction, and the human in the loop. Appl. Artif. Intell. **12**(7–8), 573–617 (1998)
13. Dautenhahn, K., Bond, A., Cañamero, L., Edmonds, B.: Socially Intelligent Agents: Creating Relationships with Computers and Robots. Springer, Cham (2002)
14. Eccles, D.W., Arsal, G.: The think aloud method: what is it and how do I use it? Qual. Res. Sport Exercise Health **9**(4), 514–531 (2017)
15. Fogg, B.J.: Creating persuasive technologies: an eight-step design process. In: Proceedings of the 4th International Conference on Persuasive Technology, pp. 1–6 (2009)
16. Fogg, B.J.: Persuasive technology: using computers to change what we think and do. Ubiquity **2002**(December), 2 (2002)
17. Fujitsu Limited: Technology for behavioral analysis Actlyzer: AI that understands, predicts, and judges like a human being. https://www.fujitsu.com/global/about/research/technology/actlyzer/. Accessed 18 Dec 2024

18. Fujitsu Limited: Fujitsu launches AI platform “Fujitsu Kozuchi”, streamlining access to AI and ML solutions to contribute to a sustainable society (2023). https://www.fujitsu.com/global/about/resources/news/press-releases/2023/0420-02.html. Accessed 18 Dec 2024
19. Gao, Y., et al.: Retrieval-augmented generation for large language models: a survey. arXiv preprint arXiv:2312.10997 (2023)
20. Gothelf, J.: Lean UX: Applying Lean Principles to Improve User Experience. O’Reilly Media Inc. (2013)
21. Hekler, E.B., et al.: Agile science: creating useful products for behavior change in the real world. Transl. Behav. Med. **6**(2), 317–328 (2016). https://doi.org/10.1007/s13142-016-0395-7
22. Hietbrink, E.A.G., et al.: A digital lifestyle coach (E-Supporter 1.0) to support people with type 2 diabetes: participatory development study. JMIR Hum Factors **10**, e40017 (2023). https://doi.org/10.2196/40017. https://humanfactors.jmir.org/2023/1/e40017
23. Hofstra, N., Petkova, B., Dullaert, W., Reniers, G., De Leeuw, S.: Assessing and facilitating warehouse safety. Saf. Sci. **105**, 134–148 (2018)
24. Isailovic, V., et al.: Compliance of head-mounted personal protective equipment by using YOLOv5 object detector. In: 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), pp. 1–5. IEEE (2021)
25. Ji, J., Krishna, R., Fei-Fei, L., Niebles, J.C.: Action genome: actions as compositions of spatio-temporal scene graphs. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10236–10247 (2020)
26. Keizer, J., Jong, N.B., Naiemi, N.A., van Gemert-Pijnen, J.: Persuading from the start: participatory development of sustainable persuasive data-driven technologies in healthcare. In: Gram-Hansen, S.B., Jonasen, T.S., Midden, C. (eds.) PERSUASIVE 2020. LNCS, vol. 12064, pp. 113–125. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-45712-9_9
27. Kihlstrom, J.F., Cantor, N.: Social intelligence. Handb. Intell. **2**, 359–379 (2000)
28. Kip, H., de Jong, N.B., Kelders, S.M., van Gemert-Pijnen, L.J.: The CeHRes Roadmap, chap. 7, pp. 103–102, 2 edn. Routledge (2024). https://doi.org/10.4324/9781003302049-9
29. Klein, C., DeRouin, R.E., Salas, E.: Uncovering workplace interpersonal skills: a review, framework, and research agenda. Int. Rev. Ind. Organ. Psychol. **2006**(21), 79–126 (2006)
30. Nelson, L.: Managing managers in occupational health and safety. Asia Pac. J. Hum. Resour. **32**(1), 13–28 (1994)
31. Niu, Y., Fan, Y., Ju, X.: Critical review on data-driven approaches for learning from accidents: comparative analysis and future research. Saf. Sci. **171**, 106381 (2024)
32. Prameswara, D., Djunaidi, Z.: Occupational health and safety in warehouse area. In: International Conference of Occupational Health and Safety (ICOHS 2017) (2018)
33. Pu, H., Yang, X., Li, J., Guo, R.: AutoRepo: a general framework for multimodal LLM-based automated construction reporting. Expert Syst. Appl. **255**, 124601 (2024)
34. Rahman, M.N.A., Zuhaidi, M.F.A.: Musculoskeletal symptoms and ergonomic hazards among material handlers in grocery retail industries. In: IOP Conference Series: Materials Science and Engineering, vol. 226. IOP Publishing (2017)
35. Reiman, T., Pietikäinen, E.: Leading indicators of system safety-monitoring and driving the organizational safety potential. Saf. Sci. **50**(10), 1993–2000 (2012)

36. RodríÂŋguez, M.C.: Íñigo Cuiñas: Definition, chap. 4 (in "Design thinking for engineering: a practical guide"), pp. 57–72. IET manufacturing series, Institution of Engineering and Technology (2023). https://doi.org/10.1049/PBME024E_ch4. https://digital-library.theiet.org/doi/abs/10.1049/PBME024E_ch4
37. Saric, S., Bab-Hadiashar, A., Hoseinnezhad, R., Hocking, I.: Analysis of forklift accident trends within Victorian industry (Australia). Saf. Sci. **60**, 176–184 (2013)
38. Schön, D.A.: The Reflective Practitioner: How Professionals Think in Action. Routledge (2017)
39. Shah, I.A., Mishra, S.: Artificial intelligence in advancing occupational health and safety: an encapsulation of developments. J. Occup. Health **66**(1), uiad017 (2024)
40. Stanford School: Design Thinking Bootleg. https://dschool.stanford.edu/resources/design-thinking-bootleg. Accessed 10 Dec 2024
41. Taype, G., Calani, M.: Extending persuasive system design frameworks: an exploratory study. In: Rocha, Á., Ferrás, C., Montenegro Marin, C.E., Medina García, V.H. (eds.) ICITS 2020. AISC, vol. 1137, pp. 35–45. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-40690-5_4
42. Tonkin, M., Vitale, J., Herse, S., Williams, M.A., Judge, W., Wang, X.: Design methodology for the UX of HRI: a field study of a commercial social robot at an airport. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, pp. 407–415 (2018)
43. Tschimmel, K.: Design thinking as an effective toolkit for innovation. In: ISPIM Conference Proceedings, p. 1. The International Society for Professional Innovation Management (ISPIM) (2012)
44. Đurđević, D., Andrejić, M., Pavlov, N.: Framework for improving warehouse safety. In: Proceedings of the 5th LOGIC Conference, pp. 304–314 (2022)
45. Vitale, J., Tonkin, M., Ojha, S., Williams, M.A., Wang, X., Judge, W.: Privacy by design in machine learning data collection: a user experience experimentation. In: 2017 AAAI Spring Symposium Series (2017)
46. Voria, G., Catolino, G., Palomba, F.: Is attention all you need? Toward a conceptual model for social awareness in large language models. In: Proceedings of the 2024 IEEE/ACM First International Conference on AI Foundation Models and Software Engineering, pp. 69–73 (2024)
47. Vukicevic, A.M., Petrovic, M.N., Knezevic, N.M., Jovanovic, K.M.: Deep learning-based recognition of unsafe acts in manufacturing industry. IEEE Access (2023)
48. Walker, R.E., Foley, J.M.: Social intelligence: its history and measurement. Psychol. Rep. **33**(3), 839–864 (1973)
49. Wang, L., Zhong, H.: LLM-SAP: large language models situational awareness-based planning. In: 2024 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), pp. 1–6. IEEE (2024)