

Lim, K. J., McGuckian, T. B., Cole, M. H., Duckworth, J., & Wilson, P. H. (2023). User experience of mixed reality applications for healthy ageing: A systematic review. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 17(4), Article 5. <https://doi.org/10.5817/CP2023-4-5>

User Experience of Mixed Reality Applications for Healthy Ageing: A Systematic Review

Kianying Joyce Lim¹, Thomas B. McGuckian¹, Michael H. Cole¹, Jonathan Duckworth², & Peter H. Wilson¹

¹ Healthy Brain and Mind Research Centre, School of Behavioural and Health Sciences, Faculty of Health Sciences, Australian Catholic University, Melbourne, Australia

² School of Design, RMIT University, Melbourne, Australia

Abstract

Mixed reality (MR) technologies are being used increasingly to support healthy ageing, but past reviews have concentrated on the efficacy of the technology. This systematic review provides a synthesis of recent experimental studies on the instrumental, emotional and non-instrumental aspects of user experience of healthy older adults in relation to MR-related applications. The review was listed on PROSPERO, utilised a modified PICOS framework, and canvassed all published work between January 2010 to July 2021 that appeared in major databases (Scopus, PubMed, CINAHL, Web of Science, and the Cochrane Library). The literature search revealed 15 eligible studies. Results indicated that all included studies measured the instrumental quality of their applications, all but two studies measured the emotional reactions triggered by gameplay, and only six studies examined participants' perception of non-instrumental quality of the applications. All included studies focused on improving a health domain such as cognitive or physical training. This suggests that the instrumental quality of the MR applications remains the focus of user experience studies, with far fewer studies examining the non-instrumental quality of the applications. Implications for game design and future research are discussed.

Keywords: mixed reality; e-games; user experience; healthy ageing; systematic review

Editorial Record

First submission received:
September 19, 2022

Revisions received:
April 6, 2023
June 1, 2023

Accepted for publication:
July 10, 2023

Editor in charge:
David Smahel

Introduction

Interactive digital technology has tremendous capacity to facilitate and promote physical activity, social participation and is seeing increasing application in healthy ageing. In particular, the use of gamified applications that incorporate Mixed Reality (MR) technologies—which includes Augmented Reality (AR) and Virtual Reality (VR) technologies—has gained interest in the health industry as a medium to enhance user engagement, motor rehabilitation and physical exercise (Cavalcanti et al., 2018). According to the World Health Organisation, healthy ageing is the process of developing and maintaining one's functional ability to enable well-being in later life (Michel & Sadana, 2017). This can include rehabilitation, as healthy ageing does not necessarily mean being free of disease. However, much of the research on the use of technology has been on rehabilitation. The focus of this review is instead on how MR technology can be used to support healthy older adults in remaining physically active and participating socially, potentially preventing the need for long-term care. Healthy older adults, for the scope

of this paper, was taken as older adults with good health who were not hospitalised, part of rehabilitation trial or diagnosed with a disorder of ageing that impacts their physical and/or cognitive ability.

The application of MR and related technology to healthy ageing has been motivated by several significant changes in population demographics: an ageing population; an increase in life expectancy, and the need to redress the progressive decline in function with older age. Furthermore, there is concern that social and health inequalities may be exacerbated by a digital divide in society, where certain population groups, amongst them older adults, have unequal access to digital health technology (Latulippe et al., 2017). Older adults face barriers to accessing technology, including technology literacy, technophobia, computer anxiety, age-related physical differences and general acceptance of technology (Di Giacomo et al., 2019; Ijaz et al., 2022; Pang & Cheng, 2023). Without equity in access, older adults may be unable to fully benefit from the technological advancement in the healthy ageing field, which points to the importance of ensuring a high user experience of MR-related applications for healthy ageing. Hence, this systematic review aims to increase the understanding of how MR technology can be used to support healthy ageing in healthy older adults.

Some of the most frequently researched technologies in aged care are MR, AR, and VR applications, all of which involve the presentation of simulated digital environments. In the literature, these terms can be difficult to define precisely but are best conceptualised along a reality-virtuality continuum that spans complete immersion in a virtual environment at one end to being fully engaged in the real environment (but with digitally augmented elements) at the other end Milgram and Kishino (1994). More recently, Skarbez et al. (2021) have revised the reality-virtuality continuum to suggest that MR technologies go beyond visual digital displays, instead incorporating any technology that enables users to perceive both virtual and real content across different senses. This expanded definition of Skarbez et al. (2021) is adopted in our review and mirrors the current state of technology, e.g., motion capture technology (Liukkonen et al., 2015; Pyae et al., 2016) and haptics, including tangible user interfaces (Boletsis & McCallum, 2016; Meza-Kubo et al., 2014).

With the advent of digital technology in healthy ageing comes a need to better understand the experiences of older adults to ensure that their interests, needs and performance capabilities are accommodated. This is done through an evaluation of user experience (UX). UX is a multi-faceted term (Hassenzahl & Tractinsky, 2006; Saariluoma & Jokinen, 2014; Thüning & Mahlke, 2007) that encompasses both task-related (or instrumental) usability and user-perceived aspects (including emotional, hedonistic or aesthetic; Hassenzahl & Tractinsky, 2006). Traditional models of usability were focused on the achievement of behavioural goals during the use of the application. Current conceptions of UX include an examination of the application's characteristics before, during and after its use (Hassenzahl & Tractinsky, 2006; Luther et al., 2020), as well as the subjective experiences of the user at each phase (Luther et al., 2020).

This review adopts the Component of User Experience (CUE) model of UX developed by Thüning and Mahlke (2007), which posits three main aspects of user experience: (i) the perception of the application's instrumental quality (e.g., usability), (ii) the emotional reactions triggered by the application (e.g., subjective feelings of the user), and (iii) the perception of its non-instrumental quality (e.g., visual aesthetics). Each component of UX can be assessed in different ways (Hassenzahl & Tractinsky, 2006; Saariluoma & Jokinen, 2014; Thüning & Mahlke, 2007). Instrumental quality of the application focuses on the ease of use of the application as perceived by the users, scheduling and implementation of the intervention protocol, and how participant restrictions such as physical or mobility limitations and safety concerns are addressed. The subjective component of user experience includes factors such as the level of immersion, sensory experience, sense of control, competency, challenge, feelings of empathy, positive and negative feelings, and behavioural involvement during and after play. Finally, the non-instrumental component covers aspects of aesthetics and subjective appraisal of the application's appeal. These are summarised in Table 1. The use of CUE model is essential to this review in that it places equal value in the instrumental quality of applications, which is what is traditionally examined in usability studies, and the emotional reactions and non-instrumental quality of application. This allows for a comprehensive assessment of the UX of MR applications used by healthy older adults.

Table 1. *Components of User Experience.*

User Perception of Instrumental Quality of Application	Emotional Reactions from Use of Application	User Perception of Non-Instrumental Quality of Application
<ul style="list-style-type: none"> - Ease of use - Duration of device use - Safety concerns - Perceived usefulness 	<ul style="list-style-type: none"> - Level of immersion - Sense of control - Sensory experience - Sense of efficacy - Sense of competency - Sense of challenge - Social engagement - Behavioural involvement - Negative affects - Positive affects 	<ul style="list-style-type: none"> - Visual aesthetics - Subjective appraisal of attractiveness/appeal of device and content - Haptic quality

While it is imperative to understand the needs and concerns of healthy older adults in the design of MR applications to maximise their benefits, few systematic reviews have focused on technology use by older adults in community and residential aged care settings. Rather, the central focus of the available reviews has been effectiveness of these applications, primarily on physical well-being and health outcomes, quality of life, motivation and social functioning (Buyl et al., 2020; Dermody et al., 2020; Miller et al., 2014). While MR technology shows promising benefits for physical health, studies included in these reviews were of low-to-moderate quality (Buyl et al., 2020; Dermody et al., 2020; Miller et al., 2014). This is due to small sample sizes, short trial duration, lack of longitudinal studies, and low quality (non-RCT) research designs. These reviews also suggested that most studies have focused on improving physical outcomes, while the impact of MR on other non-physical outcomes was typically overlooked. In recent work from 2022, reviews on this area have continued to focus on the effectiveness and outcomes of digital interventions in healthy ageing (Balki et al., 2022; Chu et al., 2022; Liu et al., 2022; Margrett et al., 2022; Suleiman-Martos et al., 2022). Notably, some of these reviews focus on field research, such as that of older adults in long-term residential care (Chu et al., 2022) and community-dwelling older adults (Suleiman-Martos et al., 2022), which is a welcomed development towards improved generalisability beyond the confines of research labs. Overall, these findings suggested that more research was required to ascertain the acceptance level and usability of MR applications by older adults.

Aspects of user experience in older adults have been addressed in two recent reviews of the literature, one on the use of VR for life enrichment (Thach et al., 2020) and the other on the feasibility of VR systems for physical activity (Miller et al., 2014). Both reviews showed that most studies were case studies with no comparators, and used a variety of outcome measures, limiting conclusions about feasibility for healthy ageing. More specifically, Thach et al. (2020) concluded that VR applications (like Oculus Rift™, Samsung Gear VR™ and Nintendo Wii™ Games) were varied in design for *enrichment purposes*, defined as experiences that support older adults with maintaining their emotional needs and/or social connections. These applications ranged from immersive VR (like head mounted display such as Oculus Rift and Samsung Gear VR) to exergames such as Wii games. The authors suggested several design and implementation considerations to enhance user experiences. The first concerned the presence of facilitators, including trainers, aged care workers, or family members, who monitor activity to ensure the VR tasks were performed correctly and safely by the older adult. Second, they noted that healthy older adults preferred more active experiences and intuitive design features to render the VR devices simple to use and “age friendly”, e.g., ensuring that the devices are comfortable (Thach et al., 2020).

The review of Miller et al. (2014) showed that available studies (up to 2,012) provided little evidence on the feasibility of home-based VR systems (mostly commercially available technologies such as Nintendo Wii games and motion capturing systems) to improve the physical health of older adults. Feasibility outcomes were reported inconsistently (e.g., recruitment information, retention rate, adherence and acceptability outcomes). The review did show, however, a strong retention and adherence rate (between 63% to 100%), based on six of 14 studies. More encouraging was the finding that unwanted side-effects were rare, including cybersickness and physical pain or over-exertion after MR use, provided that guidance and monitoring were in place, especially in the initial phase of use (Miller et al., 2014; Thach et al., 2020). Unfortunately, aside from conventional VR systems, both Miller and Thach failed to address use of (non-standard) MR applications for healthy ageing, such as bespoke solutions.

Objective

While there is growing interest in using MR and related technologies with older adults, the existing literature shows a focus on effectiveness outcomes, while overlooking usability and user experience. Even then, these reviews reveal low quality studies, a focus on conventional VR applications, and outcome measures of limited scope. Taken together, more emphasis is needed on the UX of MR-based applications for ageing in healthy older adults, especially to inform future design of these applications, larger scale evaluation studies and, ultimately, translation into viable solutions for active living.

The broad aim of the systematic review presented here was to examine the user experience (UX) of healthy older adults who enlist any MR-based technology for physical and social activity. This aim was achieved using a comprehensive three-factor model of UX, including the instrumental quality, emotional reaction and non-instrumental quality of the applications. The research questions addressed were as follows: **(RQ1)** How is user experience of MR-related technologies evaluated in recent studies involving healthy older adults? **(RQ2)** Are the currently available applications viewed as usable solutions by healthy older adults? The implications of findings for the design of future MR-based applications are discussed.

Methods

This systematic review was completed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and the PRISMA checklist (provided in Table A1) was used to ensure reporting of all relevant information (Page et al., 2021). The protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO, CRD42021266164) before the commencement of data extraction. Papers were searched, selected, evaluated, analysed, and synthesised according to the protocol described below.

Search Strategy

A multi-database systematic literature search was conducted to identify articles published between 1 January 2010 and 31 July 2021, written in English, and indexed by the SCOPUS, PUBMED, CINAHL, WEB OF SCIENCE, or COCHRANE databases. The objective and search strategy were developed based on the PICOS framework: (1) Population: older adults above 60 years of age without a diagnosis of a age-related disorder, such as dementia or Parkinson's disease; (2) Interest: any evaluation of user experience of MR applications in populations of older adults who have not been diagnosed with a disorder of ageing; (3) Comparison: comparing with a non-MR application, non-older adults population or without comparison; (4) Outcome: measures of user experience; (5) Study type: experimental and non-experimental studies.

Following this framework, relevant search descriptors were combined with the Boolean operators (OR, AND) to form the final search string as follows: ("older adults" OR elderly OR aged) AND ("virtual reality" OR "augmented reality" OR "computer aided" OR gaming OR e-games OR "mixed reality") AND (usability OR "user experience" OR feasibility).

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria are detailed below, and summarised in Table 2.

Participants

Studies involving participants aged 60 years or over were considered. Given our focus on ageing in healthy older adults, studies were excluded if they had a participant sample that included older adults who were hospitalised, evaluated as part of a rehabilitation trial or diagnosed with a disorder of ageing such as dementia or Parkinson's disease.

Interest

Studies that evaluated the user experience of MR/VR/AR and related technologies in support of healthy ageing were included in the review. As mentioned in the introduction, a MR application is one where a user can simultaneously perceive both real and virtual content and includes both AR and VR applications. AR-based applications integrate digital information layered with information about the physical environment, supplementing the real-world environment by augmenting the user experience with digitalised sensory input such as sound and visuals. Conversely, VR-based applications deliver computer-generated simulations of virtual environments with the use of immersive multimedia (L. N. Lee et al., 2019). For this review, applications termed “non-immersive VR applications” were included as these would be considered MR applications. Applications reviewed may be designed for either individual or group-based programs. Studies that only evaluated the effectiveness of the MR/VR/AR applications were excluded from the review.

Comparators

Studies were eligible for inclusion if they had either; i) a comparison between the treatment group and active or inactive control interventions such as placebo, no-treatment, standard care, treatment as usual (TAU), or a waiting list; ii) a comparison with another form of intervention (other than TAU); iii) a within-group or pre-post comparison; or iv) no comparisons, such as case studies.

Outcomes

This review considered studies that evaluated the usability and/or user-experience of MR and related applications, captured via either quantitative (e.g., user experience questionnaires) or qualitative measures (e.g., interviews).

Types of Studies

Peer-reviewed studies with experimental/controlled trials and non-experimental designs were considered in this review. Non-experimental designs include post-test only and pre-post designs. Experimental designs cover both randomised and non-randomised controlled trials using active or inactive control conditions. Studies were excluded if they were conference abstracts, theses, review papers or not written in English.

Table 2. Inclusion and Exclusion Criteria.

Criteria	Inclusion	Exclusion
Participants	Aged 60+	Older adults who were hospitalised, part of rehabilitation trial or diagnosed with a disorder of ageing
Interest	Studies that evaluated UX of MR and related technologies, including “non-immersive VR applications”	Studies that only evaluated effectiveness of MR and related applications
Comparators	Studies with either - Comparison between the treatment group and active or inactive control interventions - A comparison with another form of intervention - A within-group or pre-post comparison or - No comparisons	
Outcomes	Studies that evaluated the usability and/or user-experience of MR and related applications	
Types of studies	Peer-reviewed studies Papers written in English	Conference abstracts, theses, review papers Papers not written in English

Study Selection and Analysis

Results from the database search were included in a bibliographic database using the Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia). Duplicates were identified and removed electronically and manually. Titles and abstracts of the papers retrieved using the search strategy were first screened for potential inclusion by two evaluators (JL, TM). The full-text of potentially eligible papers were retrieved and assessed for eligibility for inclusion by two independent evaluators (JL, TM). Where there was no consensus on independent reviews, papers would be discussed and a third assessor (PW) asked to provide a review as required before a decision was made on the eligibility of the study. All discrepancies in this review were resolved by consensus.

Data Extraction and Quality Assessment

Data were extracted from the articles by one author (JL). Any uncertainty regarding data extraction was resolved by discussion among the authors. The following data were extracted: Bibliographic information (authors, title, journal, publication date), sample (participant demographics and recruitment setting), intervention (type of intervention, device(s) used, trial duration) and usability and user experience outcome measures.

The methodological reporting quality was assessed using an adapted Downs and Black (D&B) Checklist for the Assessment of Methodological Quality that was specifically modified by Wronikowska et al. (2021) for quality assessment of usability evaluation. A modified assessment was required as many of the studies that evaluated UX did not do so as part of an intervention, and hence standard quality assessment questions that pertain to an experimental design or RCT would not apply to these studies. The modified D&B Checklist has ten questions that examined the clarity of participant description, whether the sample chosen was representative of intended end-users, the inclusion of a clearly defined time period, evidence-based methods that were supported by peer-reviewed literature were chosen, that the results were clearly described and reflected the methods used.

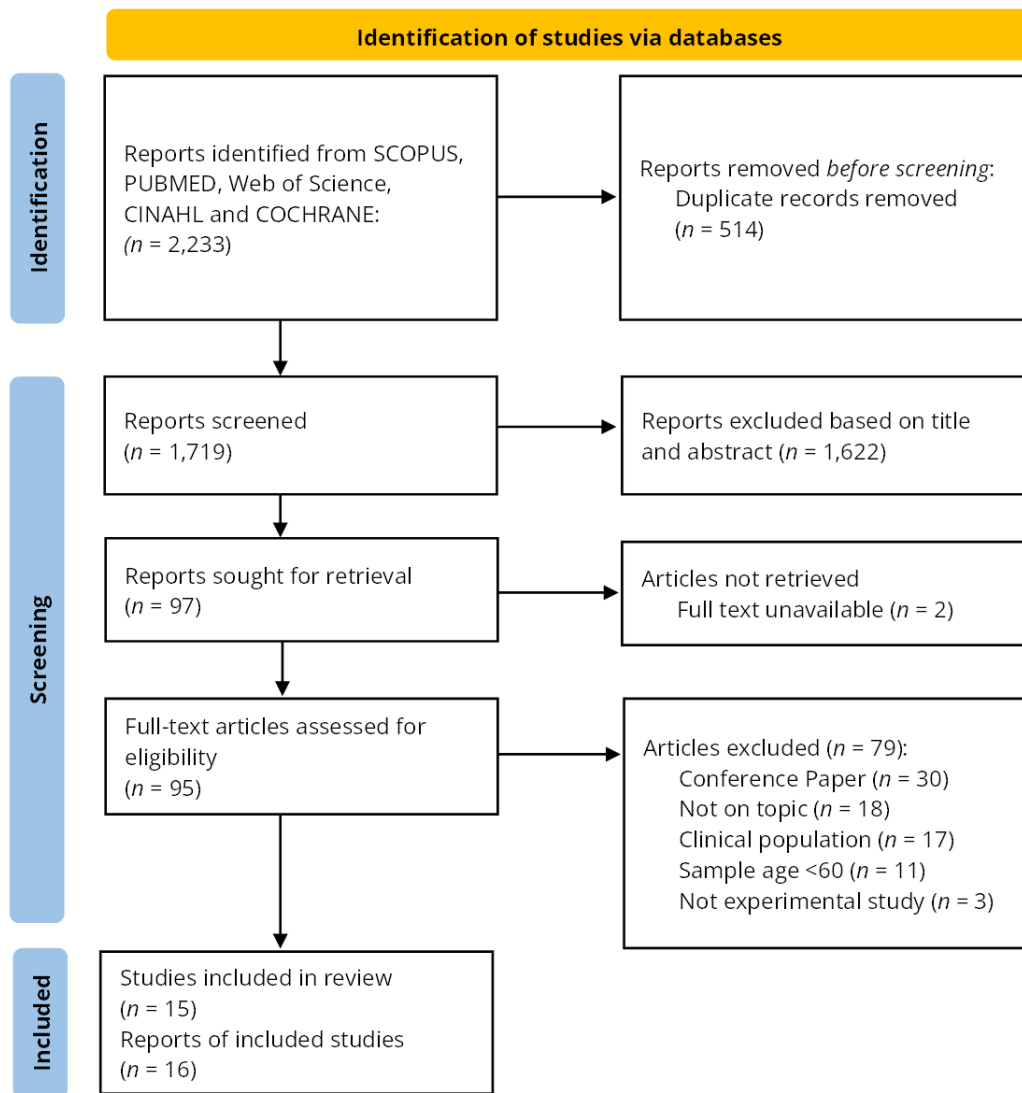
Two evaluators (JL, TM) independently assessed the quality of the selected studies. Any discrepancies were discussed, and a third evaluator (PW) consulted if no consensus was reached after the initial discussion. The quality of the included studies was coded into levels of "low", "medium" or "high" quality based on their design, conduct and analysis. High-quality studies were those that scored at or above eight out of ten points across the domains. Moderate-quality studies were those that scored a total of six or seven. Any papers that scored five and below were considered low-quality studies and were omitted from the analyses.

Results

Study Selection

The flow of studies through the systematic review process is illustrated in Figure 1. The initial search of five databases yielded 2,233 publications. After removing 514 duplicates, the titles and abstracts of the remaining 1,719 publications were screened. Of these, 97 full-text articles were assessed for eligibility. In the final phase of screening, 81 articles were excluded, of which two (Heo & Ahn, 2019; Kim et al., 2005) were due to the inability to retrieve the full-text of the articles. Emails requesting the full text of these articles were sent to the lead authors, but no reply was received. Excluded articles; i) were conference papers ($n = 30$); ii) did not have a focus on MR ($n = 18$); iii) studied a clinical population ($n = 17$); iv) involved a study sample younger than 60 years old ($n = 11$); or v) were discussion papers ($n = 3$). Of the 16 articles remaining for inclusion, two articles were found to have used the same sample (or part thereof) and, hence, were considered as one study (Pyae et al., 2016, 2017). This led to 15 studies (16 articles) being included in the final selection.

Figure 1. PRISMA Flow Chart of Database Search.



Quality Analysis

Results of the quality analysis indicated that most studies ($n = 12$) were of high-quality, with the remaining three articles of medium-quality (Meza-Kubo et al., 2014; Santos et al., 2019; Valiani et al., 2017). These studies were marked down mainly due to insufficient data reported on UX evaluation methods, participant characteristics and results. The methodological reporting quality scores for these studies are summarised in Table A2 and A3. Of the quality criteria, the included studies scored lowest on their reporting of the time period over which the study was carried out, as five studies were unclear on their trial period.

Data Synthesis

Data was extracted from the included studies and presented in two tables for ease of comparison. The first table summarises the main characteristics of the included studies (Table 3), while the second table summarises the main UX findings, as reported by the authors (Table 4).

Population Included in the Review

While the included studies all involved healthy older adults aged over 60 years, there was considerable variation in the sample sizes ($n = 5$ to $n = 236$), with seven of the included studies involving fewer than 15 participants each. Similarly, there was a large difference in the reported mean participant age (67 to 81 years). Five of the studies included the age range without reporting the mean age (Liukkonen et al., 2015; Meza-Kubo et al., 2014; Postal & Rieder, 2019; Syed-Abdul et al., 2019; Valiani et al., 2017), with Syed-Abdul et al. (2019) reporting the widest age

range of these studies (60-95 years). None of the included studies utilised random sampling. Seven of the studies recruited participants from the community, six recruited from aged care centres or homes, and two studies did not include details of their recruitment (Boletsis & McCallum, 2016; Meza-Kubo et al., 2014). One study specifically compared results between Western (Finnish) and Eastern (Japanese) cultures (Pyae et al., 2017) while four studies were conducted in an Eastern culture (Hong Kong, Korea, Singapore and Taiwan).

Types of Technologies Utilised

The included MR studies can be sorted into three main categories: studies that used a primarily haptics-based, exergame or AR/VR headset interface. Five of the studies used haptics-based applications; four papers used customised tangible devices (Boletsis & McCallum, 2016; Chen et al., 2018; Meza-Kubo et al., 2014; Wang et al., 2020), while one paper used a customised phone application (Santos et al., 2019). Six of the studies were exergames, with all using commercially available motion-detection systems such as Microsoft Kinect™, SENSO™ system, “Xtreme Reality”™ and “Extreme Motion”™ technology (N. Lee et al., 2021; Liukkonen et al., 2015; Pyae et al., 2016, 2017; Rebsamen et al., 2019; Valiani et al., 2017; Zhang et al., 2021). Two of the four AR/VR headset applications used the Kinect movement sensor in conjunction with the Oculus Rift™ (Postal & Rieder, 2019; Soares et al., 2021), one used the Vive™ VR system and one did not state the device model (Chan et al., 2020).

The Kinect system was used in four studies across exergame and AR/VR headset user interfaces and was the most commonly used commercially-available technology. All studies involving AR/VR headset technology used commercially available hardware, while three of the five haptics-based applications utilised a customised interface.

Outcomes Targeted in the Interventions

Of the included articles, six interventions focused on improving physical outcomes, five on both physical and cognitive outcomes, three primarily on cognitive outcomes, and one on cognitive outcome and sensory stimulation. All six of the interventions that focused on improving physical outcomes can be described as exergames, ranging from skiing (Pyae et al., 2016, 2017), aerobics and light exercises (Syed-Abdul et al., 2019; Valiani et al., 2017), dance (Liukkonen et al., 2015), high-intensity interval training (HIIT; Rebsamen et al., 2019) and cycling (N. Lee et al., 2021). This overlaps with the type of interventions used by studies that targeted both physical and cognitive outcomes, with three of the five studies also based on exergames (Chen et al., 2018; Postal & Rieder, 2019; Soares et al., 2021; Zhang et al., 2021). Other than exergames that are like non-virtual exercises, exergame intervention could also be like the gaming task designed by Soares et al. (2021), which required participants to virtually reach out for objects while walking on a virtual path. Differing from exergames, Santos et al. (2019) designed an AR card collection mobile game that encourages participants to venture outside of their homes for cognitive and physical stimulation.

Of the interventions that focused on cognitive outcomes, two used customised, tangible objects with gameplay like that of boardgames (Boletsis & McCallum, 2016; Meza-Kubo et al., 2014), while one used an immersive VR experience that allows users to virtually view different scenery (Chan et al., 2020). The study by Wang et al. (2020) was the only intervention that explicitly stated that it targeted sensory stimulation together with cognitive training and involved the use of a screen-based musical game that encouraged users to play a game to generate songs that sounded nostalgic to them.

User Experience

Most of the included studies ($n = 11$) utilised a quantitative measure of UX in the form of questionnaires, four studies (Boletsis & McCallum, 2016; Liukkonen et al., 2015; Pyae et al., 2016, 2017; Wang et al., 2020) reported both quantitative and qualitative measures (interviews, focus-group discussion, observations). None of the included studies used a purely qualitative measure of user experience. Of the questionnaires used, the System Usability Scale (Brooke, 1996) was used by almost half of the included studies ($n = 7$). Three of these studies used the Game Experience Questionnaire (Ijsselstein et al., 2013) in conjunction with the System Usability Scale. Details of the quantitative and qualitative measures used in each study can be found in Table 3. The UX of the included interventions can be assessed through their instrumental quality, emotional responses from the use of intervention, and non-instrumental quality, as elaborated on in the introduction.

Instrumental Quality

The trial schedule utilised by the studies was inconsistently reported, with four studies not clearly stating the number of gameplay sessions (Chen et al., 2018; Liukkonen et al., 2015; Pyae et al., 2016, 2017; Wang et al., 2020). Four studies examined the UX through more than one session of gameplay. Zhang et al. (2021) trialled their Ping Pong application once weekly for six weeks, Valiani et al. (2017) trialled their aerobics exergame application twice weekly for four weeks, Rebsamen et al. (2019) trialled their HIIT exergame thrice weekly for four weeks, and Syed-Abdul et al. (2019) trialled their VR exergame twice weekly for six weeks. Studies with multiple trials were similarly all exergame-based, which may explain the methodological choice of having multiple trial sessions, with a consensus that the health impact of exercise is best shown with consistent engagement over time.

The duration of gameplay also varied widely, from 30-seconds of gameplay (Postal & Rieder, 2019) to around 30-minutes (Boletsis & McCallum, 2016; Rebsamen et al., 2019; Valiani et al., 2017). Most studies utilised approximately 20 to 30 minutes of gameplay. The longest duration of trial was by Santos et al. (2019), who examined a mobile game application that measured participants over two weeks of free-play on the app.

Where reported, included papers described mild to no adverse reactions such as cybersickness. Chan et al. (2020) was the only paper to report its drop-out rate and suggested that drop-out rate was low when the MR application was compared with traditional gameplay. Several studies suggested the need for better communication with older adults, including a need for a clearer explanation of device use (Boletsis & McCallum, 2016; Santos et al., 2019) and more visual and aural feedback to increase understanding and immersion (Postal & Rieder, 2019). There is also a need to customise the level of gameplay to a comfortable level for older adults. This includes adjusting the game pace (Liukkonen et al., 2015), difficulty level (Soares et al., 2021) and using settings that were more familiar to older participants (Liukkonen et al., 2015; Syed-Abdul et al., 2019).

Emotional Reaction

Examining the emotional reactions triggered by gameplay, all included studies suggested that their application was perceived to be enjoyable. However, there were differing views on whether enjoyment of gameplay or perceived usefulness was a bigger motivating factor. Syed-Abdul et al. (2019) found that the enjoyment level perceived by older adults increased intention of use more than its perceived usefulness, and Postal and Rieder (2019) found that participants' interest in gameplay was independent of their task performance during the trial, suggesting that enjoyment was an important motivator. In contrast, Chen et al. (2018) found that participants' acceptance of the game was more dependent on the perceived usefulness by the gamer than perceived enjoyment.

Few papers reported social presence in gameplay, and those that did, reported a low level of social presence and involvement with other users (Liukkonen et al., 2015; Santos et al., 2019).

Non-Instrumental Quality

Six studies reported non-instrumental factors such as users' appraisal of the devices' aesthetics (Chen et al., 2018; Postal & Rieder, 2019; Pyae et al., 2016, 2017; Rebsamen et al., 2019; Santos et al., 2019; Wang et al., 2020). These studies indicated that when non-instrumental factors are examined, older adults expressed a need for intuitive visual and aural feedback to increase their engagement (Postal & Rieder, 2019; Rebsamen et al., 2019), yet without too much information shown on the game graphics (Wang et al., 2020). Participants also indicated a preference for games simulated in real-world environment and are related to sports and real-world activities (Pyae et al., 2016, 2017). Furthermore, participants indicated that the output quality impacted their enjoyment of (Chen et al., 2018) and engagement with the game (Rebsamen et al., 2019).

Table 3. Main Characteristics of Included Studies.

Author/Year	Sample (age; gender)	Country/Recruitment Setting	Physical Hardware	Targeted Outcomes	Description of MR Application	Trial Schedule Trial: no. of session Gameplay: X minutes	UX assessment measure(s) <i>Qualitative</i> <i>Quantitative</i>
<i>Type of User Interface: Haptics Based MR Applications</i>							
Boletsis & McCallum (2016)	<i>n</i> = 5 (mean = 67.6, range 61–75; gender not reported)	not reported	Tangible cubes customised for intervention, tablet PC on base stand	Cognitive Training	Six mini AR games for cognitive training	Trial(s): 1 trial + 1 main session Gameplay: 20–25 min (trial); 25–30 min (actual)	<i>Quantitative:</i> GEQ; SUS; <i>Qualitative:</i> Semi-structured user interviews
Chen et al. (2018)	<i>n</i> = 39 (mean = 79; male = 15, female = 24)	Taiwan/Community	Floor projector, interactive tables, touch screen computer	Cognitive and Physical Training	Three physical games and two cognitive exergames	Not stated	<i>Quantitative:</i> TAM questionnaire
Meza-Kubo et al. (2014)	<i>n</i> = 6 (range = 65–75; gender not reported) and caregivers <i>n</i> = 6 (35–55; gender not reported)	not reported	Customised tangible interface	Cognitive Training	Tangible surface intervention designed for cognitive stimulation	Trial(s): 1 Gameplay: 10 min MR + 10 min augmented traditional board games (comparison)	<i>Quantitative:</i> Customised TAM based questionnaire
Santos et al. (2019)	<i>n</i> = 12 (mean = 75; male = 5, female = 7)	Community	Smartphone Application	Cognitive and Physical Training	AR card collection game for cognitive and physical stimulation	Trial(s): N.A. Gameplay: 2 weeks free-play	<i>Quantitative</i> GEQ
Wang et al. (2020)	<i>n</i> = 30, (mean = 81; male = 8, female = 22)	Aged care	Customised tangible interface interactive device	Cognitive Training and Sensory Stimulation	Screen-based game designed for cognitive and sensory stimulation	Trial(s): not reported Gameplay: 15 min Total duration: 40 min	<i>Quantitative:</i> 2-part questionnaire based on SUS and Questionnaire for User Interface Satisfaction (QUIS) <i>Qualitative:</i> User interviews, expert interviews; In-field observations

Type of User Interface: Exergame Applications

N. Lee et al. (2021)	<i>n</i> = 5; (mean = 69; male = 3, female = 2)	Korea/Community	Visual display, stationary recumbent cycle	Physical Training	Immersive cycling training system encouraging physical activity	Trial(s): 1 Gameplay: 20 min	<i>Quantitative:</i> SUS, SSQ
Liukkonen et al. (2015)	<i>n</i> = 19 (range = 61–83; male = 9, female = 10)	Finland/Community (Rural and Urban)	“Extreme Motion” motion detector	Physical Training	One large muscle exercise game, one body maintenance game	Trial(s): not reported Gameplay: under 15 min	<i>Quantitative:</i> SUS, GEQ <i>Qualitative:</i> Observation records, participant interviews
Pyae et al. (2016, 2017)	Finnish Trial: <i>n</i> = 21 (mean = 71; male = 8, female = 13); Japanese trial: <i>n</i> = 24 (mean = 72; male = 12, female = 12)	Finland and Japan/ Finnish Study: elderly service home; Japanese Study: health promotion centre	“Xtreme Reality” Technology	Physical Training	Skiing exergame for encouraging physical activity	Trial(s): not reported Finnish study: Gameplay: 5–10 min Japanese study: Gameplay: up to 20 min (including demo)	(Both studies) <i>Quantitative:</i> GEQ, SUS <i>Qualitative:</i> post-gameplay interview
Valiani et al. (2017)	<i>n</i> = 12 (range = 76.3–84.7; male = 2, female = 10)	Florida/Nursing home	Jintronix hardware, TV screen, Microsoft Kinect motion tracking camera	Physical Training	Virtual aerobic, strength and balance exercise designed for physical training	Trial(s): 2/wk x 4wks Gameplay: 30 min	<i>Quantitative:</i> Customised feasibility and acceptability questionnaire; data from exergame platform
Zhang et al. (2021)	<i>n</i> = 33 (mean = 70.9; male = 9, female = 24)	Singapore/Community	Kinect V2 movement sensor, Full HD TV, laptop	Cognitive and Physical Training	Ping Pong exergame for cognitive inhibition training	Trial(s): 1/wk x 6wks Gameplay: 15–20 min	<i>Quantitative:</i> SUS
Rebsamen et al. (2019)	<i>n</i> = 12 (mean = 72.3; male = 2, female = 10)	Switzerland/Community	SENSO exercise system (pressure sensitive platform)	Physical Training	Set of nine high intensity interval exergames for physical training	Trial(s): 3/wk x 4 wks Gameplay: 19–35 min	<i>Quantitative:</i> TAM, SUS <i>Qualitative:</i> Participants’ general statements recorded during gameplay and when answering questionnaire

Type of User Interface: AR/VR Headsets Applications

Chan et al. (2020)	<i>n</i> = 236 (mean = 73.85; male = 56, female = 180)	Hong Kong/Elder care centres	VR device, model not reported	Cognitive Training	VR cognitive stimulation activity	Trial(s): 1 Gameplay: 20–25 min	<i>Quantitative:</i> SSQ, Positive and Negative Affect Score (PANAS)
Postal & Rieder (2019)	<i>n</i> = 20 (range 60–81; male = 4, female = 16)	Elder care centre	Oculus Rift, Kinect motion sensor Comparison: Smart 3D TV	Cognitive and Physical Training	Exergame designed to improve motor and cognitive abilities	Trial(s): 1 Gameplay: 30 seconds per device	<i>Quantitative:</i> Customised usability questionnaire
Soares et al. (2021)	<i>n</i> = 10 (mean = 70; male = 3, female = 7) comparison group: <i>n</i> = 11 (mean = 22; male = 3, female = 8)	São Paulo/Community	Oculus Rift, Microsoft Kinect Sensor	Cognitive and Physical Training	Three virtual reaching tasks completed while walking on a virtual path	Trial (s): 1 Gameplay: 30 min	<i>Quantitative</i> SSQ, satisfaction Questionnaire (customised), Use of Technology Questionnaire, International Physical Activity Questionnaire
Syed-Abdul et al. (2019)	<i>n</i> = 30 (range 60–95; male = 6, female = 24)	Aged care centre	Vive HTC VR system	Physical Training	Exergame with nine applications targeting physical stimulation	Trial(s): 2/wk x 6 wks Gameplay: 15 min	<i>Quantitative:</i> Adapted TAM Scale

Note. 2/wk x 6 wks = 2 sessions per week for 6 weeks

Table 4. Main UX Findings in Included Studies.

Author/Year	Main Findings—Instrumental UX	Main Findings—Emotional UX	Main Findings—Non-Instrumental UX
<i>Type of User Interface: Haptics Based MR Applications</i>			
Boletsis & McCallum (2016)	Average system usability score (70.5; SUS) Users experienced loss of depth perception when interacting in the real world while watching the output on the tablet's screen There is a need to simplify interaction techniques to prevent confusion between actual and AR markers	Overall high scores of Positive affect, Immersion and Challenge (iGEQ) Some mini-games had high values of Negative affect and Tension and low Flow (iGEQ) System errors contribute to negative feelings of confusion, uncertainty and tension	N.A.
Chen et al. (2018)	Participant's prior experience in computer use had little to no impact on perceived usefulness (TAM2) Perceived ease of use was not related to usage behaviour, potentially because participants had limited prior experience in technology	Perceived playfulness and usefulness were positively related to intention to use and usage behaviour	Satisfaction with output quality (game interface and appearance) was correlated with perceived playfulness (enjoyment of application) Output quality was correlated with usage behaviour for cognitive games, suggesting its importance in designing cognitive games
Meza-Kubo et al. (2014)	All participants perceived the application as easy to use and useful	High perceived enjoyment of the application, with higher enjoyment of entertainment game versus cognitive game, potentially due to more familiarity with the entertainment activity Overall level of anxiety was low for all participants, but older adult participants had slightly higher level of anxiety than the younger relatives who accompanied them on the trials Participants reported high levels of intention to use.	N.A.
Santos et al. (2019)	25% attrition rate (3 out of 12) after one week due to health reasons and difficulties with device use Participants reported difficulties in understanding the game controls Game rules and goals were easily understood	Participants liked the level of challenge and reported high levels of engagement, satisfaction, motivation to play and enjoyment Participants reported weak sense of social presence and involvement with other people (GEQ)	Participants liked the game's visual style but disliked the game music
Wang et al. (2020)	Participants reported a high level of perceived usability, with sufficient level of perceived ease of use and perceived usefulness Participants and expert interviews indicated that the use of device can be improved with simpler and more intuitive controls and clearer instructions on usage	Participants reported finding the game interesting and experienced joy while gaming Participants indicated that using the device can improve their relationships with others and increase their willingness to join public activities Participants worried about forgetting how to play the game	Participants reported finding the device attractive Participants reported satisfaction with the colour brilliance and richness of game, with clearly perceived audio and visual stimuli Participants indicated a need to reduce the amount of information shown on the game graphics.

Type of User Interface: Exergame Applications

N. Lee et al. (2021)	High system usability scores (94.60; SUS), indicating high levels of perceived acceptability and ease of use Low levels of cybersickness with minor symptoms of nausea or disorientation (SSQ)	N.A.	N.A.
Liukkonen et al. (2015)	Game A (large muscle exercise game) failed to reach minimum acceptable usability (SUS :58.29) while Game B (body maintenance exercise) was at average level (SUS: 79.44) Item analysis suggested Game A was perceived as too complex and cumbersome; Authors also noted that the pace of Game B was calmer Interviews suggest most participants felt they could play the games without help, but some participants felt they may need help with home use due to being unaccustomed to the technology	Overall, more positive emotional experience in Game B (higher scores of Positive affect, Competence, Immersion, Flow and lower Negative affect; GEQ) Both games had low levels of psychological involvement with the other participants (i.e., social presence)	N.A.
Pyae et al. (2016, 2017)	Both Finnish and Japanese participants perceived the game positively in terms of perceived usability Comparison of results suggested greater acceptability by the Japanese participants Both groups did not experience disorientation after gameplay Participants from both groups indicated a need for visual cues due to their poorer eyesight and lack of experiences in digital gameplay and indicated a preference for voice-based instructions in their own language Qualitative interviews indicated that some participants had experienced minor symptoms of cybersickness	Both Finnish and Japanese participants reported similarly high levels of positive affect in-game and post-game Japanese participants reported significantly higher level of perceived negative affect, tension and tiredness than Finnish participants although overall perceived negative affect was still low The Japanese groups perceived the game as more challenging than Finnish participants Both groups reported average levels of Flow, Immersion and Competence, indicating moderate levels of interest in gameplay Qualitative interviews indicated that low achievement in the first instance of gameplay would decrease motivation to continue playing in the future Novice game players reported a preference for simple gameplay over cognitively challenging gameplay Participants reported a fear of getting injured while playing the game, indicating a need to ensure that gameplay is designed to minimise physical risks	Both groups of participants indicated a preference for games simulated in real-world environment and that relates to sports and real-world activities Participants reported enjoying controller-free gameplay
Valiani et al. (2017)	Participants were able to complete 86.3% of the total sessions with acceptable level in the quality of movements	Participants reported a global appreciation score of 91.7%, indicating high level of acceptance. Difficulty level of the intervention was slightly lower than the expected minimum level.	N.A.

Zhang et al. (2021)	Participants with and without prior experience in electronic device usage both reported an above-average level of system usability with well-understood game objectives (SUS > 68)	N.A.	N.A.
Rebsamen et al. (2019)	No reported adverse events and 8% participant attrition rate 91% protocol adherence rate High levels of acceptance rate (TAM) and satisfaction (SUS: 93.5) and perceived usefulness Qualitative findings suggest gameplay and usage of device were easily understood	Participants found games to be highly enjoyable Qualitative findings suggest combining physical training with games was a motivating factor Qualitative findings also suggest gameplay can be made more challenging to increase motivation	Qualitative findings suggest gameplay can be more engaging, such as with the inclusion of music Qualitative findings found that participants wanted more visual and auditive feedback while playing to increase engagement
<i>Type of User Interface: AR/VR Headset Applications</i>			
Chan et al. (2020)	Minimal adverse events from VR activity, only 1.4 % of participants had severe simulator sickness (SSQ) Only minimal supervision required for VR use after initial training No significant difference in drop-out rates between VR (28.7%) and paper-and-pencil activities (19.8%)	Mostly comparable feelings of negative and positive affects between VR and paper-and-pencil activities (PANAS) Participants were significantly more excited with paper-and-pencil activity	N.A.
Postal & Rieder (2019)	No significant difference in perceived ease of use, ease of task execution, clarity of procedures and level of comfort between immersive headset and 3D TV device Observation of the sessions indicated that participants were uncomfortable when using 3D devices due to unfamiliarity with technology 40% of participants expressed that having a training session to familiarise themselves with the gameplay would increase their performance 20% of participants had persistent difficulties in spatial orientation (e.g., not perceiving the need to open arms wider to accomplish task goal)	Participants' interest in the gameplay was independent of their task performance, with users perceiving the application as interesting and beneficial	Observations of gameplay indicate that more intuitive visual and aural feedbacks is required to aid older adult users in understanding tasks
Soares et al. (2021)	Low feasibility, with only 20% of older adult participants completing all three tasks, compared to 81.8% of young adult participants All older adult participants had mild cybersickness symptoms, with 8 participants stopping use of the system after first signs of cybersickness Execution of the tasks without supervision is not recommended due to the need for a safety harness	Older adult participants were more motivated to practice the game and recommend it to others when compared with young adult participants Most participants reported that they would not play the game at home due to difficulties in setting up the device and safety harness	N.A.

Technical issues in the system could have contributed to high incidence of cybersickness

Syed-Abdul
et al. (2019)

Perceived usefulness and ease of use predicted users' intention to use

Perceived enjoyment is a better predictor of intention to use than perceived usefulness, suggesting that enjoyment is an important motivator

N.A.

Discussion

This systematic literature review provides an overview of the user experience of healthy older adults with MR applications that are designed to enhance their physical activity and social interaction. The results provide evidence to support use of MR technology in this population across a range of applications that include haptics, exergame and AR/VR headset interfaces. Although the quality assessment indicates that most studies were of high-quality, the diverse range of MR applications, UX measures and health outcomes made it difficult to compare results across studies. Results showed that instrumental quality and emotional reactions to MR-based applications were commonly measured, while non-instrumental factors were not. Overall, the instrumental and emotional quality of the investigated MR applications were high. The discussion below considers this pattern of results and its implications for the design of MR applications for healthy ageing.

Do the Included Studies Measure All Aspects of UX?

A main research objective of this paper was to find out how the user experience of MR-related technologies was evaluated in recent studies involving healthy older adults. While all 15 studies measured the instrumental quality of their applications, and all but two measured emotional reactions triggered by gameplay (N. Lee et al., 2021; Zhang et al., 2021), only six studies examined non-instrumental qualities of the application (Chen et al., 2018; Postal & Rieder, 2019; Pyae et al., 2016, 2017; Rebsamen et al., 2019; Santos et al., 2019; Wang et al., 2020). The limited information on non-instrumental qualities makes it challenging to understand how the aesthetics of the applications may affect participants' level of engagement and usage behaviour.

When comparing UX measurements between different types of MR applications, we observed that studies on AR/VR headset applications had largely neglected measurement of non-instrumental UX. This finding was surprising given that the aesthetics of AR/VR headset applications, where game design focuses on the visuals of the immersive virtual environment, would likely impact the UX of a MR application. Many of the AR/VR headset applications were pre-designed games that did not appear to be customised specifically (or easily) for older adults (Soares et al., 2021; Syed-Abdul et al., 2019). For example, the same virtual reaching task was tested on both young and older adults (Soares et al., 2021). In comparison, when customised interfaces were used, measurements of non-instrumental UX were more likely, exemplified by Chen et al. (2018) and Wang et al. (2020) in the haptics-based application category. Both studies showed that game attractiveness, audio and visual output quality were associated with user satisfaction.

How Was the User Experience of MR Application Perceived by Healthy Older Adults ?

Based on the findings of studies that measured non-instrumental quality of the MR applications, games that mirror real-life activities and sports, such as simulation of skiing, cycling and walking outdoors (N. Lee et al., 2021; Pyae et al., 2016; Soares et al., 2021), were reported to appeal to the users. While it has been suggested that older adults may prefer such games (Pyae et al., 2017), the precise reason for this preference remains unclear; five of the six studies that assessed non-instrumental quality did not examine the players' perception on the internal content of gameplay. Instead, these studies focused more on the quality of visual and auditory outputs from the devices per se--enjoyment of gameplay was associated with the richness and quality of visual and auditory outputs (Chen et al., 2018; Pyae et al., 2016, 2017; Rebsamen et al., 2019; Wang et al., 2020). It appears that the familiarity of game content appealed more to the older adult users: with familiarity comes a sense of comfort that may enhance both engagement and motivation (Zhang et al., 2019). The positive impact of familiarity is even more important for older adults who have less technology experience (Hwang et al., 2011), although further research is required to confirm this hypothesis.

In relation to instrumental quality, few studies reported adverse impacts of the MR applications on their users, suggesting MR applications are generally safe to use. There were, however, some isolated instances where gameplay induced feelings of unease and/or mild nausea. For example, challenges were evident in the study by Soares et al. (2021), where 80% of the older adults could not complete all three game tasks due to mild cybersickness symptoms. In addition, the included papers have generally indicated a high level of perceived system usability, perceived usefulness, and perceived ease of use, further supporting their use with older adult populations. An exception is the study of Liukkonen et al. (2015), which reported that one of their games was

perceived as too complex and cumbersome by the older adult users, such as confusion about the game objectives and finding the game speed to be too fast, leading to a low usability score. Together, although adverse events are uncommon, developers should continue to carefully test their applications to ensure game design accommodates the physical and cognitive capabilities of older adults and their past experience with MR technologies, relative to younger adults.

In relation to the emotional quality of applications, an intriguing connection is evident between game complexity on the one hand, and levels of engagement, motivation and performance on the other for older adults. In game design, the level of challenge is known to impact user engagement. In younger student populations, Hamari et al., (2016) found that the challenge of the game can increase engagement when learning such games. With older participants, the study of Pyae and colleagues (2017) showed that older adults may perceive less challenging games as more enjoyable, and that games that were too difficult at first use may reduce their motivation to keep playing (Pyae et al., 2017). Games that are too complex may also hinder the older participants' ability to participate in gameplay (see Liukkonen et al., 2015, and Soares et al., 2021). This may be particularly evident when there are differences in computer literacy within the older adult population, with some just beginning to use digital technologies and others having kept up with the latest technologies (Tyler et al., 2020). By comparison, two studies have indicated that the amount of user-interaction complexity in their game design motivated gameplay (Rebsamen et al., 2019; Santos et al., 2019). These mixed results prompt further research on how game complexity may impact the emotional quality of the MR application.

Specifically, future research should investigate how the level of challenge affects motivation for gameplay, and how best to balance the need to scale game complexity for training purposes with the need to keep older adults motivated and using the devices. A Dutch study on the motivations of older adults participating in VR training suggested that both enjoyment and challenge, conceptualised as the drive to perform, were intrinsic motivators (De Vries et al., 2018). The concept of challenge was associated with receiving positive visual feedback, so that participants were aware when they perform well, and with a game speed that is fast enough to increase challenge without overstressing older adult participants (De Vries et al., 2018). By comparison, Subramanian et al. (2020) showed that older adults were intrinsically motivated by the enjoyment of exergame play and not by the level of challenge, unlike younger adults (Subramanian et al., 2020). Both studies did agree, however, that enjoyment and perceived effectiveness of the game were important motivators for older adults (De Vries et al., 2018; Subramanian et al., 2020). Future research can also look into the use of dynamic difficulty adjustment (DDA) in game design (Ang & Mitchell, 2017; Cantwell et al., 2012). DDA is where the complexity level of the game is balanced against each user's performance and skill. This also allows players of different levels to play together. There is little available data about how DDA works with healthy older adult users, but some studies have looked into using DDA in rehabilitation exergames for older adults (Pezzerà & Borghese, 2020; Smeddinck et al., 2013). In this review, a third of the studies were based in an Eastern culture (Chan et al., 2020; Chen et al., 2018; N. Lee et al., 2021; Pyae et al. 2016, 2017; Zhang et al., 2021), while 27% were based in a Western culture and the remainder (40%) provided insufficient detail to be certain of their setting. There is insufficient information on cultural differences in the included studies to form a commentary on the cultural impact on UX, with only Pyae et al. (2017) comparing participants across different cultures, it would be beneficial to understand how the environment of the participants may impact their experience of using MR applications. Future research should investigate cultural and socio-economic factors that may influence UX of MR applications for healthy ageing.

What Are the Considerations for Future Game Design?

The lack of information of non-instrumental quality of the MR applications for healthy ageing is concerning because MR technology is being enlisted to provide both an entertaining and attractive way for older adults to stay active. By overlooking aesthetics in game design, even simple games that are user-friendly may not attract and engage users in the longer term, thereby reducing their potential health benefits. De Schutter and Abeele (2015) have boldly proposed that focusing primarily on the benefits and accessibility of gameplay diminishes the intrinsic value of gaming. More useful and meaningful is a game design framework that takes into account the aesthetic and ergonomic preferences of the players themselves (De Schutter & Abeele, 2015). Indeed, when included as game co-designers, older adult participants have been shown to identify aesthetics as essential elements of fun and attractive gameplay (Havukainen et al., 2020). Taken together, it is recommended that future game design take into serious consideration the aesthetics of the application.

Instrumental quality of MR applications may be enhanced in the future by ensuring clear user instructions, improving both the user's sense of control when interacting with the virtual environment (Witmer & Singer, 1998) and intuitiveness of game control. Studies that used participant interviews found that improving the clarity of instruction and intuitive design of user interfaces and controls can enhance the usability of applications (Boletsis & McCallum, 2016; Liukkonen et al., 2015). Several studies used trained facilitators on-site to explain game process and answer any questions the participants had during game play (Chan et al., 2020; Meza-Kubo et al., 2014). The earlier systematic review by Thach et al. (2020) also showed that UX was improved when facilitators were around to provide guidance on the usage of the applications. As well, the clarity of instructions can also be improved by having a guided game tutorial session before actual gameplay (Pyae et al., 2016), which provides an opportunity for users to learn and familiarise themselves with the interface (Liukkonen et al., 2015), and is especially important when older adult users may be unfamiliar with technology. Liukkonen et al. (2015) also suggested that a longer play session would give older adult users more time to learn the game, which may help them to relax and immerse themselves in gameplay.

With respect to the fidelity of study design, user experience was only examined across one session of gameplay in most studies, with limited data on user experience after extensive use of the MR application. It was also common for studies to use a small sample size, impacting the methodological quality of studies and generalisability of results. Eleven out of 15 studies utilised only quantitative measures of UX, missing an opportunity for deeper insight into the older adults' experience of gameplay by including qualitative measures of UX. All studies reviewed had a focus on improving aspects of health like cognition or physical activity. This focus may reflect the assumption that ageing is a period of decline in cognitive and physical abilities (De Schutter & Abeelee, 2015), instead of a life-stage like any other that can be enjoyed and offer new opportunities for game play and socialising. To redress this assumption, future research should investigate the non-instrumental qualities of user experience. Games that are designed for play (both competitive and cooperative) rather than for improved health could be a positive addition to the discourse on MR applications that promote healthy ageing.

In general, it would benefit game development if older adults themselves were enlisted as co-designers—after all, they are the ones who best understand their experience of game use (De Schutter & Abeelee, 2015; Havukainen et al., 2020). The concept of co-designing was not explored by any of the included studies in this review. Involvement of older adults as co-designers from the initial conceptualisation phase of game design will help game developers scale the level of challenge according to their needs and preference, design an interface that is attractive to them, and ensure that game controls are intuitive and easily understood.

Limitations

Wide variation in the type of MR application and use of less rigorous research designs has precluded a meta-analysis of results. We anticipate that heightened interest in MR technology will see a much larger body of work in the near future, supporting meta-analysis.

Our systematic review excluded conference papers because they more commonly use low quality research designs, inadequately report information, are difficult to access full papers, and lack peer review. Studies published in languages other than English were also omitted due to challenges in professionally translating the papers. It is acknowledged that this exclusion may introduce bias in the results. In effect, a number of more recent (non-journal and non-English) papers were excluded from this review.

Strengths

The strengths of this review lie in providing a comprehensive look at how studies in the area of MR-applications for healthy ageing have analysed the different aspects of UX—the instrumental, emotional and non-instrumental components. This provides researchers and practitioners with a better understanding of how technology may be perceived by older adults, going beyond perceived functional benefits, and also pointing to gaps in the research and design of MR-related applications. The studies included in this review were carried out across different countries and cultures, and showed the vibrancy of the research in this area.

Conclusion

Results of this systematic review have indicated that the instrumental quality of the MR applications remains the focus of UX studies, with far fewer studies examining their non-instrumental quality. Encouragingly, these studies show promising results on conventional aspects of instrumental quality and emotional experience. The neglect of non-instrumental quality is, however, surprising in view of the fact that such aspects of UX are most likely to impact the experience and behaviour of older adults, especially when using audio-visual-based MR applications, such as AR/VR headset applications. Co-design frameworks are a useful approach to game design that may help redress this limitation.

Future research needs to examine the multi-faceted nature of UX, especially its non-instrumental qualities, and track user experience across a longer period of use to confirm its usability and acceptance level. In doing so, research should consider the use of UX questionnaires that examine all three components of the CUE model of UX (Thüring & Mahlke, 2007), such as the meCUE questionnaire (Minge et al., 2017).

Conflict of Interest

The authors have no conflict of interests to declare.

Authors' Contribution

Kianying Joyce Lim: conceptualization, data curation, methodology, writing—original draft, writing—review & editing. **Thomas B. McGuckian:** conceptualization, data curation, methodology, writing—original draft, writing—review & editing. **Michael H. Cole:** conceptualization, methodology, supervision, writing—original draft, writing—review & editing. **Jonathan Duckworth:** conceptualization, writing—review & editing. **Peter H. Wilson:** conceptualization, supervision, writing—original draft, writing—review & editing.

Acknowledgement

While completing this research the lead author was supported by an Australian Government Research Training Program (RTP) Scholarship.

References

- Ang, D., & Mitchell, A. (2017). Comparing effects of dynamic difficulty adjustment systems on video game. In CHI PLAY '17: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (pp. 317–327). ACM. <https://doi.org/10.1145/3116595.3116623>
- Balki, E., Hayes, N., & Holland, C. (2022). Effectiveness of technology interventions in addressing social isolation, connectedness, and loneliness in older adults: Systematic umbrella review. *JMIR Aging*, 5(4), Article e40125. <https://doi.org/10.2196/40125>
- Boletsis, C., & McCallum, S. (2016). Augmented reality cubes for cognitive gaming: Preliminary usability and game experience testing. *International Journal of Serious Games*, 3(1), 3–18. <https://doi.org/10.17083/ijsg.v3i1.106>
- Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry* (pp. 189–194). CRC Press. <https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke>
- Buyl, R., Beogo, I., Fobelets, M., Deletroz, C., Van Landuyt, P., Dequanter, S., Gorus, E., Bourbonnais, A., Giguère, A., Lechasseur, K., & Gagnon, M.-P. (2020). E-health interventions for healthy aging: A systematic review. *Systematic Reviews*, 9(1), Article 128. <https://doi.org/10.1186/s13643-020-01385-8>
- Cantwell, D., Broin, D. O., Palmer, R., & Doyle, G. (2012). Motivating elderly people to exercise using a social collaborative exergame with adaptive difficulty. In P. Felicia (Ed.), *Proceedings of the 6th European Conference on Games Based Learning* (pp. 615–619). Academic Publishing International Limited Reading.

- Cavalcanti, V. C., de Santana, M. I., Da Gama, A. E. F., & Correia, W. F. M. (2018). Usability assessments for augmented reality motor rehabilitation solutions: A systematic review. *International Journal of Computer Games Technology*, 2018, Article 5387896. <https://doi.org/10.1155/2018/5387896>
- Chan, J. Y. C., Chan, T. K., Wong, M. P. F., Cheung, R. S. M., Yiu, K. K. L., & Tsoi, K. K. F. (2020). Effects of virtual reality on moods in community older adults. A multicenter randomized controlled trial. *International Journal of Geriatric Psychiatry*, 35(8), 926–933. <https://doi.org/10.1002/gps.5314>
- Chen, C.-K., Tsai, T.-H., Lin, Y.-C., Lin, C.-C., Hsu, S.-C., Chung, C.-Y., Pei, Y.-C., & Wong, A. M. K. (2018). Acceptance of different design exergames in elders. *PLoS One*, 13(7), Article e0200185. <https://doi.org/10.1371/journal.pone.0200185>
- Chu, C. H., Quan, A. M. L., Souter, A., Krishnagopal, A., & Biss, R. K. (2022). Effects of exergaming on physical and cognitive outcomes of older adults living in long-term care homes: A systematic review. *Gerontology*, 68(9), 1044–1060. <https://doi.org/10.1159/000521832>
- De Schutter, B., & Abeele, V. V. (2015). Towards a gerontoludic manifesto. *Anthropology & Aging*, 36(2), 112–120. <https://doi.org/10.5195/aa.2015.104>
- De Vries, A. W., van Dieën, J. H., van den Abeele, V., & Verschueren, S. M. P. (2018). Understanding motivations and player experiences of older adults in virtual reality training. *Games for Health Journal*, 7(6), 369–376. <https://doi.org/10.1089/g4h.2018.0008>
- Dermody, G., Whitehead, L., Wilson, G., & Glass, C. (2020). The role of virtual reality in improving health outcomes for community-dwelling older adults: Systematic review. *Journal of Medical Internet Research*, 22(6), Article e17331. <https://doi.org/10.2196/17331>
- Di Giacomo, D., Ranieri, J., D'Amico, M., Guerra, F., & Passafiume, D. (2019). Psychological barriers to digital living in older adults: Computer anxiety as predictive mechanism for technophobia. *Behavioral Sciences*, 9(9), Article 96. <https://doi.org/10.3390/bs9090096>
- Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, 54, 170–179. <https://doi.org/10.1016/j.chb.2015.07.045>
- Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, 25(2), 91–97. <https://doi.org/10.1080/01449290500330331>
- Havukainen, M., Laine, T. H., Martikainen, T., & Sutinen, E. (2020). A case study on co-designing digital games with older adults and children: Game elements, assets, and challenges. *The Computer Games Journal*, 9(2), 163–188. <https://doi.org/10.1007/s40869-020-00100-w>
- Heo, S.-P., & Ahn, D. H. (2019). Group-type rehabilitation system for virtual reality-based dementia prevention and cognitive training. *Journal of Advanced Research in Dynamical and Control Systems*, 11(12 Special Issue), 941–947. <https://doi.org/10.5373/JARDCS/V11SP12/20193297>
- Hwang, M.-Y., Hong, J.-C., Hao, Y.-w., & Jong, J.-T. (2011). Elders' usability, dependability, and flow experiences on embodied interactive video games. *Educational Gerontology*, 37(8), 715–731. <https://doi.org/10.1080/03601271003723636>
- Ijaz, K., Tran, T. T. M., Kocaballi, A. B., Calvo, R. A., Berkovsky, S., & Ahmadpour, N. (2022). Design considerations for immersive virtual reality applications for older adults: A scoping review. *Multimodal Technologies and Interaction*, 6(7), Article 60. <https://doi.org/10.3390/mti6070060>
- IJsselsteijn, W. A., de Kort, Y. A. W., & Poels, K. (2013). *The game experience questionnaire*. Technische Universiteit Eindhoven. https://pure.tue.nl/ws/files/21666907/Game_Experience_Questionnaire_English.pdf
- Kim, M.-Y., Lee, K.-S., Choi, J.-S., Kim, H.-B., & Park, C.-I. (2005). Effectiveness of cognitive training based on virtual reality for the elderly. *Journal of the Korean Academy of Rehabilitation Medicine*, 29(4), 424–433. <https://www.e-arm.org/journal/view.php?number=1676>
- Latulippe, K., Hamel, C., & Giroux, D. (2017). Social health inequalities and eHealth: A literature review with qualitative synthesis of theoretical and empirical studies. *Journal of Medical Internet Research*, 19(4), Article e136. <https://doi.org/10.2196/jmir.6731>

- Lee, L. N., Kim, M. J., & Hwang, W. J. (2019). Potential of augmented reality and virtual reality technologies to promote wellbeing in older adults. *Applied Sciences*, 9(17), Article 3556. <https://doi.org/10.3390/app9173556>
- Lee, N., Choi, W., & Lee, S. (2021). Development of an 360-degree virtual reality video-based immersive cycle training system for physical enhancement in older adults: A feasibility study. *BMC Geriatrics*, 21(1), Article 325. <https://doi.org/10.1186/s12877-021-02263-1>
- Liu, M., Zhou, K., Chen, Y., Zhou, L., Bao, D., & Zhou, J. (2022). Is virtual reality training more effective than traditional physical training on balance and functional mobility in healthy older adults? A systematic review and meta-analysis. *Frontiers in Human Neuroscience*, 16, Article 125. <https://doi.org/10.3389/fnhum.2022.843481>
- Liukkonen, T. N., Pitkäkangas, P., Heinonen, T., Makila, T., Raitoharju, R., & Ahtosalo, H. (2015). Motion tracking exergames for elderly users. *IADIS International Journal on Computer Science and Information Systems*, 10(2), 52–64. <http://www.iadisportal.org/ijcsis/papers/2015180204.pdf>
- Luther, L., Tiberius, V., & Brem, A. (2020). User experience (UX) in business, management, and psychology: A bibliometric mapping of the current state of research. *Multimodal Technologies and Interaction*, 4(2), Article 18. <https://doi.org/10.3390/mti4020018>
- Margrett, J. A., Ouerson, K. M., Gilbert, S. B., Phillips, L. A., & Charness, N. (2022). Older adults' use of extended reality: A systematic review. *Frontiers in Virtual Reality*, 2, Article 176. <https://doi.org/10.3389/frvir.2021.760064>
- Meza-Kubo, V., Morán, A. L., & Rodríguez, M. D. (2014). Bridging the gap between illiterate older adults and cognitive stimulation technologies through pervasive computing. *Universal Access in the Information Society*, 13(1), 33–44. <https://doi.org/10.1007/s10209-013-0294-3>
- Michel, J.-P., & Sadana, R. (2017). "Healthy aging" concepts and measures. *Journal of the American Medical Directors Association*, 18(6), 460–464. <https://doi.org/10.1016/j.jamda.2017.03.008>
- Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, 77(12), 1321–1329. https://www.researchgate.net/publication/231514051_A_Taxonomy_of_Mixed_Reality_Visual_Displays
- Miller, K. J., Adair, B. S., Pearce, A. J., Said, C. M., Ozanne, E., & Morris, M. M. (2014). Effectiveness and feasibility of virtual reality and gaming system use at home by older adults for enabling physical activity to improve health-related domains: A systematic review. *Age and Ageing*, 43(2), 188–195. <https://doi.org/10.1093/ageing/aft194>
- Minge, M., Thüring, M., Wagner, I., & Kuhr, C. V. (2017). The meCUE questionnaire: A modular tool for measuring user experience. In *Advances in ergonomics modeling, usability & special populations* (pp. 115–128). Springer. https://link.springer.com/chapter/10.1007/978-3-319-41685-4_11
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal*, 372, Article n71. <https://doi.org/doi:10.1136/bmj.n71>
- Pang, W. Y. J., & Cheng, L. (2023). Acceptance of gamified virtual reality environments by older adults. *Educational Gerontology*. Advance online publication. <https://doi.org/10.1080/03601277.2023.2166262>
- Pezzera, M., & Borghese, N. A. (2020). Dynamic difficulty adjustment in exer-games for rehabilitation: A mixed approach. In *2020 IEEE 8th International Conference on Serious Games and Applications for Health*. IEEE. <https://doi.org/10.1109/SeGAH49190.2020.9201871>
- Postal, M. R. B., & Rieder, R. (2019). A usability evaluation experiment of 3D user interfaces for elderly: A pilot study. *Revista Brasileira De Computacao Aplicada*, 11(3), 28–38. <https://doi.org/10.5335/rbca.v11i3.9477>
- Pyae, A., Joelsson, T. N., Saarenpää, T., Mika, L., Kattimeri, C., Pitkäkangas, P., Granholm, P., & Smed, J. (2017). Lessons learned from two usability studies of digital skiing game with elderly people in Finland and Japan. *International Journal of Serious Games*, 4(4), 37–52. <https://doi.org/10.17083/ijsg.v4i4.183>
- Pyae, A., Liukkonen, T. N., Saarenpää, T., Luimula, M., Granholm, P., & Smed, J. (2016). When Japanese elderly people play a Finnish physical exercise game: A usability study. *Journal of Usability Studies*, 11(4), 131–152. <http://uxpajournal.org/japanese-elderly-finnish-exercise-game-usability/>

- Rebsamen, S., Knols, R. H., Pfister, P. B., & de Bruin, E. D. (2019). Exergame-driven high-intensity interval training in untrained community dwelling older adults: A formative one group quasi-experimental feasibility trial. *Frontiers in Physiology*, 10(8), Article 1019. <https://doi.org/10.3389/fphys.2019.01019>
- Saariluoma, P., & Jokinen, J. P. P. (2014). Emotional dimensions of user experience: A user psychological analysis. *International Journal of Human-Computer Interaction*, 30(4), 303–320. <https://doi.org/10.1080/10447318.2013.858460>
- Santos, L. H. O., Okamoto, K., Hiragi, S., Yamamoto, G., Sugiyama, O., Aoyama, T., & Kuroda, T. (2019). Pervasive game design to evaluate social interaction effects on levels of physical activity among older adults. *Journal of Rehabilitation and Assistive Technologies Engineering*, 6, 1–8. <https://doi.org/10.1177/2055668319844443>
- Skarbez, R., Smith, M., & Whitton, M. C. (2021). Revisiting Milgram and Kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, 2, Article 27. <https://doi.org/10.3389/frvir.2021.647997>
- Smeddinck, J. D., Siegel, S., & Herrlich, M. (2013). Adaptive difficulty in exergames for Parkinson's disease patients. In *GI '13: Proceedings of Graphics Interface 2013* (pp. 141–148). Canadian Information Processing Society. <https://dl.acm.org/doi/10.5555/2532129.2532154>
- Soares, B. C., Bacha, J. M. R., Mello, D. D., Moretto, E. G., Fonseca, T., Vieira, K. S., de Lima, A. F., Lange, B., Torriani-Pasin, C., de Deus Lopes, R., & Pompeu, J. E. (2021). Immersive virtual tasks with motor and cognitive components: A feasibility study with young and older adults. *Journal of Aging and Physical Activity*, 29(3), 400–411. <https://doi.org/10.1123/JAPA.2019-0491>
- Subramanian, S., Dahl, Y., Skjæret Maroni, N., Vereijken, B., & Svanæs, D. (2020). Assessing motivational differences between young and older adults when playing an exergame. *Games for Health Journal*, 9(1), 24–30. <https://doi.org/10.1089/g4h.2019.0082>
- Suleiman-Martos, N., García-Lara, R., Albendín-García, L., Romero-Béjar, J. L., Cañadas-De La Fuente, G. A., Monsalve-Reyes, C., & Gomez-Urquiza, J. L. (2022). Effects of active video games on physical function in independent community-dwelling older adults: A systematic review and meta-analysis. *Journal of Advanced Nursing*, 78(5), 1228–1244. <https://doi.org/10.1111/jan.15138>
- Syed-Abdul, S., Malwade, S., Nursetyo, A. A., Sood, M., Bhatia, M., Barsasella, D., Liu, M. F., Chang, C. C., Srinivasan, K., Raja, M., & Li, Y. C. J. (2019). Virtual reality among the elderly: A usefulness and acceptance study from Taiwan. *BMC Geriatrics*, 19(1), Article 223. <https://doi.org/10.1186/s12877-019-1218-8>
- Thach, K. S., Lederman, R., & Waycott, J. (2020). How older adults respond to the use of virtual reality for enrichment: A systematic review. In *32nd Australian Conference on Human-Computer Interaction* (pp. 303–313). ACM. <https://doi.org/10.1145/3441000.3441003>
- Thüring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human-technology interaction. *International Journal of Psychology*, 42(4), 253–264. <https://doi.org/10.1080/00207590701396674>
- Tyler, M. A., De George-Walker, L., & Simic, V. (2020). Motivation matters: Older adults and information communication technologies. *Studies in the Education of Adults*, 52(2), 175–194. <https://doi.org/10.1080/02660830.2020.1731058>
- Valiani, V., Lauzé, M., Martel, D., Pahor, M., Manini, T. M., Anton, S., & Aubertin-Leheudre, M. (2017). A new adaptive home-based exercise technology among older adults living in nursing home: A pilot study on feasibility, acceptability and physical performance. *Journal of Nutrition, Health and Aging*, 21(7), 819–824. <https://doi.org/10.1007/s12603-016-0820-0>
- Wang, C.-M., Tseng, S.-M., & Huang, C.-S. (2020). Design of an interactive nostalgic amusement device with user-friendly tangible interfaces for improving the health of older adults. *Healthcare*, 8(2), Article 179. <https://doi.org/10.3390/healthcare8020179>
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3), 225–240. <https://direct.mit.edu/pvar/article-abstract/7/3/225/92643/Measuring-Presence-in-Virtual-Environments-A?redirectedFrom=fulltext>
- Wronikowska, M. W., Malycha, J., Morgan, L. J., Westgate, V., Petrinic, T., Young, J. D., & Watkinson, P. J. (2021). Systematic review of applied usability metrics within usability evaluation methods for hospital electronic

healthcare record systems: Metrics and evaluation methods for eHealth systems. *Journal of Evaluation in Clinical Practice*, 27(6), 1403–1416. <https://doi.org/10.1111/jep.13582>

Zhang, H., Miao, C., Wu, Q., Tao, X., & Shen, Z. (2019). The effect of familiarity on older adults' engagement in exergames. In *HCI 2019: Human-Computer Interaction International Conference* (pp. 277–288). Springer. https://doi.org/10.1007/978-3-030-22015-0_22

Zhang, H., Shen, Z. Q., Liu, S. Y., Yuan, D. Z., & Miao, C. Y. (2021). Ping pong: An exergame for cognitive inhibition training. *International Journal of Human-Computer Interaction*, 37(12), 1104–1115. <https://doi.org/10.1080/10447318.2020.1870826>

Appendix

Table A1. *PRISMA Checklist.*

Section and Topic	Item #	Checklist item	Location where item is reported
<i>TITLE</i>			
Title	1	Identify the report as a systematic review.	1
<i>ABSTRACT</i>			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	2
<i>INTRODUCTION</i>			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	3–6
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	6–7
<i>METHODS</i>			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	8–9
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	7–8
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	7–8
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	9–10
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	10–12
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g., for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	10–12
	10b	List and define all other variables for which data were sought (e.g., participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	10
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	12
Effect measures	12	Specify for each outcome the effect measure(s) (e.g., risk ratio, mean difference) used in the synthesis or presentation of results.	N.A.
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g., tabulating the study intervention characteristics and comparing against the planned groups for each synthesis—item #5).	11–12
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	11–12
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	11–12
	13d	Describe any methods used to synthesize results and provide a rationale	11–12

		for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g., subgroup analysis, meta-regression).	N.A.
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesised results.	N.A.
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	N.A.
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	N.A.
<i>RESULTS</i>			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	12
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	12–13
Study characteristics	17	Cite each included study and present its characteristics.	17–24
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	S2
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g., confidence/credible interval), ideally using structured tables or plots.	N.A.
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	N.A.
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g., confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	N.A.
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	N.A.
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesised results.	N.A.
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N.A.
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	N.A.
<i>DISCUSSION</i>			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	25–29
	23b	Discuss any limitations of the evidence included in the review.	30
	23c	Discuss any limitations of the review processes used.	30
	23d	Discuss implications of the results for practice, policy, and future research.	30–31
<i>OTHER INFORMATION</i>			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	7
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	7
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N.A.

Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	1
Competing interests	26	Declare any competing interests of review authors.	1
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	N.A.

Note. From: Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372. <https://doi.org/10.1136/bmj.n71>

Table A2. *Quality Analysis Using Adapted Downs and Black Checklist by Wronikowska et al. (2021).*

	Questions	Give 1 point if:
1	Is the hypothesis and aim/objective of the study clearly described?	Hypothesis and aim/objectives all clearly described
2	Are the main usability evaluation methods described in the Introduction or Methods section?	Usability evaluation methods clearly described in Introduction or Methods section
3	Are all the usability evaluation methods referenced?	Referenced the method if the method is not novel or described the method clearly if the method is novel.
4	Are the characteristics of the participants participating in the usability evaluation method/s clearly described?	The number of participants, age, gender, inclusion & exclusion criteria are all clearly described
5	Are the participants participating in the usability evaluation method/s representative of the intended eventual users?	The participant profile should fit intended end users stated in the introduction (e.g., does it fit the demographics intended to be examined?)
6	Is the time period described over which the study was carried out?	Time period described (including intervention period if applicable and duration of sessions)
7	Do the results accurately reflect the methods described?	The authors do exactly what they said they would do
8	Are all the studied usability evaluation method results described (numerically or graphically)?	Usability evaluation method results clearly described
9	Are the usability evaluation methods measured using standardised measurement tools?	A standardised measurement was used and described (e.g., the use of standardised set of questionnaire or semi-structured interview questions across all participants)
10	Do the authors apply appropriate qualitative/quantitative statistics (specific to usability/user experience evaluation methods)?	Application and sufficient reporting of statistics in case of quantitative methods (e.g., reporting of central tendency and variation), narrative analysis for qualitative methods

Table A3. Results of Quality Analysis.

Paper/Criteria	1	2	3	4	5	6	7	8	9	10	TOTAL	Quality Category
Meza-Kubo et al., 2014	0	0	1	0	0	1	1	1	1	1	6	Moderate
Santos et al., 2019	1	0	1	0	1	1	1	1	1	0	7	Moderate
Valiani et al., 2017	1	1	0	1	1	1	1	1	0	0	7	Moderate
Postal & Rieder, 2019	1	1	1	1	1	1	1	1	0	0	8	High
Boletsis and McCallum, 2016	1	1	1	0	0	1	1	1	1	1	8	High
Chen et al., 2018	1	1	1	1	1	0	1	1	1	1	9	High
Wang et al., 2020	1	1	1	0	1	1	1	1	1	1	9	High
Pyae et al., 2016, 2017	1	1	1	1	1	0	1	1	1	1	9	High
Chan et al., 2020	1	1	1	1	1	0	1	1	1	1	9	High
Liukkonen et al., 2015	1	1	1	1	1	0	1	1	1	1	9	High
Soares et al., 2021	1	1	1	1	1	0	1	1	1	1	9	High
Lee et al., 2021	1	1	1	1	1	1	1	1	1	1	10	High
Syed-Abdul et al., 2019	1	1	1	1	1	1	1	1	1	1	10	High
Rebsamen et al., 2019	1	1	1	1	1	1	1	1	1	1	10	High
Zhang et al., 2021	1	1	1	1	1	1	1	1	1	1	10	High
TOTAL	14	13	14	11	13	10	15	15	13	12		

About Authors

Kianying Joyce Lim is a doctoral candidate in the Faculty of Health Sciences, Australian Catholic University and a practising psychologist with a background in educational and developmental psychology. Her research focuses on social good, including finding ways to harness technology for health and well-being.

<https://orcid.org/0000-0003-2124-9166>

Thomas B. McGuckian is a human movement scientist with a background in both psychology and exercise science. His research focusses on the use of technology to assess human movement in natural environments, human and environment interaction, and the development of motor skill over the lifespan. He also has expertise in psychometric assessment of technology and research synthesis and meta-synthesis.

<https://orcid.org/0000-0002-5490-0042>

Michael H. Cole is an Associate Professor at Australian Catholic University with an extensive background in the fields of biomechanics and neuroscience and specific expertise with high-technology systems used for the assessment of balance, mobility, and neuromuscular function in healthy and neurologically impaired populations.

<https://orcid.org/0000-0003-1817-7528>

Jonathan Duckworth is an Associate Professor in Digital Design and co-director of CiART (Creative interventions, Art and Rehabilitative Technology), School of Design at Royal Melbourne Institute of Technology. Dr Duckworth has established a strong reputation for his interdisciplinary practice-based design research that forges synergies between interaction design, digital media art, health science, disability and game technology. His research relates to design innovation within acquired brain injury rehabilitation, the arts and human computer interaction technology field.

<https://orcid.org/0000-0002-8021-3860>

Peter H. Wilson is a Professor of Developmental Psychology in the School of Behavioural and Health Sciences, co-Director of Healthy Brain and Mind Research Centre, and program lead for the Development and Disability over the Lifespan Program. Prof. Wilson leads an internationally-renowned program of work on motor development and rehabilitation, is an International Research Fellow of Radboud University in Nijmegen and former President of the International Society for Research on Developmental Coordination Disorder (DCD).

<https://orcid.org/0000-0003-3747-0287>

✉ Correspondence to

Joyce Lim, Australian Catholic University, The Daniel Mannix Building, Level 5, 17 - 29 Young St, Fitzroy, 3065, Australia, kianyingjoyce.lim@myacu.edu.au

© Author(s). The articles in *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* are open access articles licensed under the terms of the [Creative Commons BY-SA 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/) which permits unrestricted use, distribution and reproduction in any medium, provided the work is properly cited and that any derivatives are shared under the same license.