2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) | 979-8-3503-2838-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISMAR59233.2023.00118

Augmented Reality Rehabilitative and Exercise Games (ARREGs): A Systematic Review and Future Considerations

Cassidy R. Nelson[1]
Virginia Tech

Joseph L. Gabbard[2]
Virginia Tech

Abstract

Augmented Reality (AR) and exergames have been trending areas of interest in healthcare spaces for rehabilitation and exercise. This work reviews 25 papers featuring AR rehabilitative/exercise games and paints a picture of the literature landscape. The included results span twelve years, with the oldest paper published in 2010 and the most recent work in 2022. More specifically, this work contributes a bank of representative ARREGs and a synthesis of measurement strategies for player perceptions of Augmented Reality Rehabilitative and Exercise Game (ARREG) experiences, the elements that comprise the exergame experience, the intended use cases of ARREGs, whether participants are actually representative users, the utilized devices and AR modalities, the measures used to capture rehabilitative success, and the measures used to capture participant perceptions. Informed by the literature body, our most significant contribution is nine considerations for future ARREG development.

Keywords: Augmented Reality, Exergame, Rehabilitation, Serious Game, Systematic Literature Review

Index Terms: • General and reference ~ Document types ~ Surveys and overviews • Human-centered computing ~ Human computer interaction (HCI) • Social and professional topics ~ User characteristics

1 Introduction

The intersection of well-being, health, and eudemonia is a grand challenge facing humans living and interacting within "technology-augmented environments" today [1]. This is due to the advancement of both technology and medicine, yielding new ways to promote healthy lives and eudemonia (need fulfillment and pleasure) [1], [2]. One of these outlined mechanisms is via new forms of therapy like exergames.

Serious games in the context of health have been studied from as early as 1988 [3] and are a currently trending area of interest for health-intervention research [4] in countless arenas like amblyopia [5], obesity [6], Parkinson's disease [7], stroke [8], and more. This is due to their evidenced promise to afford greater socialization [9]–[11], increased motivation [4], [12], heightened engagement [7], [13], increased energy expenditure [14], enhanced strength and endurance [15], and to trigger positive experiences [4]. Games in the health arena include exercise, rehab, distraction therapy, training on health information, cognitive rehab, mental health treatment, and more [16]. This work focuses on exergames, or games used for exercise and rehabilitation [1], [16], [17].

[1]email: cassidynelson@vt.edu
[2]email: jgabbard@vt.edu

In conjunction, extended reality (XR) technologies, comprised chiefly of augmented reality (AR) and virtual reality (VR), have also penetrated this research space due to their ability to lower injury risk by simulating dangerous scenarios [4], their ability to enhance immersion [4], their ever-improving affordability [18], their capacity to provide synchronous feedback on user movements [19], their ability to provide interventions without a therapist present [20], and the capacity to afford at-home intervention [21].

Beyond research interest, there has also been a complementary exploding consumer interest in exergames (Dance Dance Revolution, Ring Fit Adventure, and more), and XR (Meta Quest, mobile phones, etc.). This trend is also true for XR exergames like Beat Saber, PokémonGo, and Superhot. The coalescence of these two areas, XR for health and exergames, leverages each other's benefits to create a more effective experience.

The primary goal of XR exergames is to motivate players to move [22] and push themselves further [23] than they would without the intervention. XR exergames used for rehab are also used to maintain player engagement and motivation during rote, repetitive tasks [7][24][25]. This is crucial as patients "often abandon rehabilitation because… [it is] boring and repetitive," [25].

Because the primary goal of these exergames is to motivate players, it is vital to evaluate games for their actualized ability to do this. Thus, this work seeks to capture the realized participant perception of AR rehabilitative/exercise games (ARREGs) and their leveraged design metaphors. This work provides a synthesis of the ARREG literature body regarding participant outcomes, an understanding of whether the participant body is representative, and the reported justifications for AR and games in rehab. Further, this work reports the utilized game and rehab elements, extra elements that enhance the ARREG experience, and the leveraged AR modality and devices. Informed by these components, we provide nine considerations for future ARREG development.

While VR is defined as total immersion in a synthetic world [26], AR is defined as the combination of real and virtual worlds in a synchronous, interactive way [27] that enhances user perception of reality [20]. Although both AR and VR can foster profound user experiences with unique pros and cons, this review focuses on AR, specifically mobile AR, head-mounted displays (both optical see-through and video passthrough), and spatial AR (SAR) as long as the experience includes visual elements and reality is augmented in the relevant context. AR allows for more unique design metaphors (handheld, head-mounted, projection-based), and thus a more comprehensive range of exergame possibilities. Further, patients achieve better rehabilitation outcomes if they are highly engaged in participation [25], and AR is less likely to induce motion sickness [4]. This is crucial as motion sickness has been found to reduce motivation [4], [28], and interaction accuracy [4].

Interaction accuracy is particularly imperative in the ARREG context and a focus of much literature. This is because a core area of ARREG research is synchronous system feedback and asynchronous therapeutic monitoring by doctors and therapists. AR allows for synchronous feedback and corrections of user

979-8-3503-2838-7/23/$31.00 ©2023 IEEE
DOI 10.1109/ISMAR59233.2023.00118

movement that are situated in reality [20]. This, supported by contactless tracking to reduce movement hindrances [24], fosters more natural interactions [29] than those in VR. These unobstructed, natural, reality-based interactions are more enjoyable and motivating [29], safer [20], and critical for accurate progress evaluation and translatable exercise movements, i.e., movements that patients will do in their daily activities.

Within the niche context of ARREGs, there is a wide range of possible foci. Some works focus on developing new hardware to facilitate ARREG experiences, some center on creating computer vision algorithms to capture player movements, while others evaluate user perceptions of the game experiences. This work will focus on the latter.

There are a few related systematic reviews of interest. Junaini et al. reviewed AR and VR games for occupational safety and health training [30]. While meaningful for understanding trends for occupational safety and opportunities to prevent injury, it does not provide guidance for rehabilitation once an injury has occurred. Mubin et al. start to explicitly bridge the transition from clinic to home-based rehab settings [31]. However, they do not cover participant demographics, game elements, or user perceptions in their review.

Cavalcanti et al. evaluate the literature for all usability assessments utilized to examine AR rehab solutions [32]. Their work similarly captures the clinical use of the rehab solutions and categorizes the strategies used by the included works to ascertain participant usability. However, their work only includes four game-related papers. Further, our work expands beyond usability evaluations to capture all evaluation considerations by any AR rehab game.

Yim & Graham completed a review on games for exercise motivation [33]. From this, they were able to distill some potential requirements of successful exercise games for eliciting player motivation. However, this work only captured three AR games, two of which had no participants, and the third was Augmented Virtuality, or AV [33]. Milgram et al. define AV as "completely graphic display environments… to which some amount of 'reality' has been added," [26]. E.g. unlike AR, which overlays digital information into the real world, AV brings physical elements into a virtual space.

Gil et al., completed a review on AR for physical therapy but did not include any games [34]. However, they did note that game-based rehabilitation would be of interest in the context of AR [34]. Yu et al. review applications of XR games in healthcare broadly. They sought to capture evaluation techniques and applications of XR from a wider healthcare perspective [35]. Their review is the closest to ours in that they capture, in broad strokes, the primary game task (fishing, diving, etc.) and the measurements used to evaluate participant perception. However, they do not distinguish between VR and AR games and we posit that there are meaningful differences/advantages to AR that should be explored. Further, our review provides a more detailed breakdown of game elements within the included broad game tasks.

Thus, our review separates itself as a more detailed evaluation of AR-specific games, realized participant demographics, the broad evaluation methods utilized in this space, and the game elements that create the ARREG experience.

2 METHODS

2.1 Protocol

We conducted a systematic literature review guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) [36]. It should be noted that the PRISMA framework had to be adapted for our review because PRISMA is intended initially as a medical study framework for quantitative meta-analyses, and this review offers a qualitative systematic review. We chose to use PRISMA, despite this need for adaptation, due to its high prevalence in related works[34], [37]–[41] and to try and furnish a standard by which to compare this review to other reviews [42] and improve replicability. For transparency, the following PRISMA elements guide the structure of this review: Title, Abstract, Introduction (rationale, objectives), Methods (information sources, eligibility criteria, search study selection, data collection process, data items, synthesis of results), Results (study selection, study characteristics, results of individual studies), and Discussion (Summary of evidence, limitations, conclusions).

(Included PRISMA items 1-4, 6-11, 14, 17-18, 20, 24-26)

2.2 Information Sources

The utilized databases were PubMed, IEEEXplore, and JMIR. For each database, we used the following search string:

"augmented reality" AND exergame OR "augmented reality" AND rehab* AND gam* OR "augmented reality" AND exercise AND gam**

The search was completed on April 14, 2023 and updated on June 01, 2023. PubMed was used because it indexes more than 30 million citations for medical literature, yielding 56 results with our search string. (*of note, the PubMed wildcard entries were game* due to a $\geq$4 letter root requirement*). IEEEXplore indexes papers from the Computational Intelligence and Games (CIG) conference, the IEEE Transactions on Games, VS-Games, the Serious Games and Applications for Health (SeGAH) conference, and more, yielding 184 relevant results. Finally, JMIR was used as they host several journals, including JMIR Serious Games, JMIR Rehabilitation and Assistive Technologies, JMIR Human Factors, and many more. JMIR yielded 430 relevant search results. A detailed breakdown of our inclusion flow process can be seen in Figure 1.

2.3 Eligibility Criteria & Findings

Included articles adhered to these criteria:

- The paper focused on an application (not a concept paper),
- That application was a game (with game elements, not just a task labeled as a 'game'),
- That game was furnished with visual augmented reality,
- The game's core underlying goal was to facilitate physical exercise or rehabilitation (not cognitive),
- The ARREG was evaluated with participants,
- Participants were asked about their perceptions (no papers just testing the effectiveness of their algorithm),
- Games were described in some level of detail,
- The full paper is available,
- The paper is not a duplicate,
- The paper is in English

Games labeled as either rehab or exercise are both included in this review because they are highly enmeshed. Rehab involves exercise, and some exercise can be denoted as rehab. We did not see a meaningful way or reason to parse them. Finally, many researchers use the terms 'gamification' and 'serious games' interchangeably [17]; thus, this work included both labels as both seek to promote engagement through game-based interactions. Gamification is the addition of game elements to an existing rehabilitation structure or exercise (adding points and high scores to the number of kick motions done) to improve engagement [43]. In contrast, serious games are defined as an entirely new rehabilitative experience provided via a complete game [43], [44] (creating a competitive game experience where players must kick a virtual ball and score against the other player [11]).

2.4 Captured Study Components

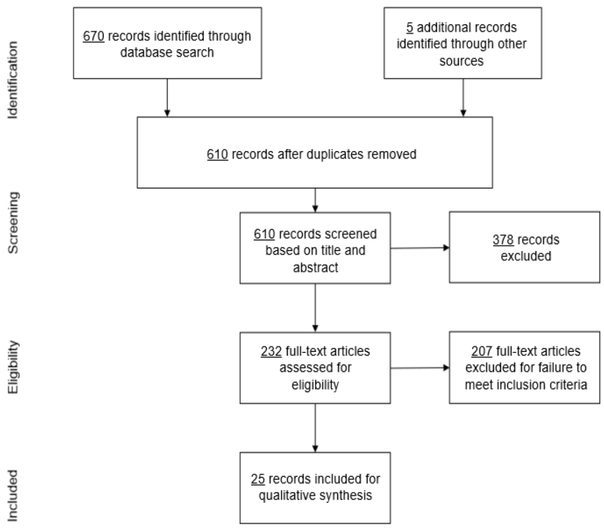

Figure 1: Inclusion and Distillation Flow Diagram

This work reviews existing ARREG applications across clinical domains to capture and synthesize the field's landscape, particularly the participants' demographics, elements of the games, rehabilitative elements, user perception, and the AR modalities.

In pursuance of this goal, this review captures:

- The game's treatment purpose,
- game elements,
- rehab/exercise elements,
- exergame elements,
- extra elements,
- target exercise/rehab demographic,
- realized participant demographics,
- citation demographics,
- AR modality (or type)
- devices used to afford the experience,
- performance/rehab measures,
- perception measures,
- and study outcomes.

3 Results & Discussion

3.1 Search Results

The aforementioned parameters resulted in 670 articles for consideration. Five additional records (totaling 675) were identified through the evaluation process of the 675 original articles, typically in lists of example-related works. 65 duplicates were excluded leaving 610. 378 records were excluded based on their title and abstract, leaving 232 articles for a more extensive, full-text review. Some works had to be eliminated only because they did not describe their game in any detail, they did not have any participants, or participants were not asked for their perception of the experience. This more extensive review eliminated 207 articles, leaving 25 for final synthesis: [4]–[7], [10], [11], [16], [18], [20], [22]–[25], [29], [45]–[55].

The purposes of the games were as follows: Twelve papers focused on specific clinical rehabilitation, nine on general exercise, and four on evaluation. The primary goal of evaluation papers was to understand how to evaluate patient progress in rehab gains or to motivate patients to thoroughly engage with the evaluation process so a complete picture of their physical capabilities could be captured.

Our included works feature four nested groups. I.e., there are sets of papers that are related/extensions of existing games. The groupings are as follows:

- [20], [45]
- [7], [23], [24], [47]
- [49], [51]
- [25], [52]

Nine papers covered more than one game, opting instead to focus on a battery of games. In total, there were 46 games featured in the review. 35 of these games are unique, i.e., only appearing once. We considered excluding these duplicate games from our analysis as one could argue that they mildly inflate the unique number of game elements used across ARREGs. However, we chose not to exclude them because each work has unique participant data. Further, these games are being iterated over time (thus changing) or being used to answer different questions. Individual game elements will be defined in section 3.4.1.

It should be noted that while 25 PokémonGo papers were found in our initial pull, none of them were included in our final analysis. We feel this is important to mention because at its height, more than 44 million people played PokémonGo [56] and, similarly, it captured the attention of both AR game researchers and exergame researchers. This research interest was driven by the game being advertised as an AR game for exercise. Unfortunately, the reality of the actual game's use is more nuanced.

Most of PokémonGo is spent looking at a fully, virtually synthetic version of a map of the player's city. We posit that this component is augmented virtuality (AV) as a completely virtual world has been augmented with real-world elements. The truly AR component of PokémonGo is an optional element only relevant when catching a Pokémon. Players navigate the world via this AV environment to find Pokémon. Once one is found, a player can opt into "AR mode," where a Pokémon will appear somewhere in the real world. The problem is that this AR mode makes the game more challenging and physically fatiguing, so quickly after release, many players turned off this function and played the game entirely in AV, or quit playing entirely due (in part) to being disappointed by the AR experience [57].

We examined the papers focusing on PokémonGo. Unfortunately, none of the papers explicitly evaluated AR usage. Therefore it is unknown whether their users leveraged the optional AR component of the game. Thus, the relationship the AR component had (if any) with any reported outcomes is unclear. Due to this, these PokémonGo papers were not included in our deeper analysis. However, we will broadly note that PokémonGo was found responsible for significant increases in physical activity over a period of 30 days [58], positive behavioral changes [59], and greater environmental exploration [60], thus showing meaningful promise for at least exergames.

3.2 Justification for AR Games (Unique Benefits)

We captured the reported unique benefits of AR and games in the context of exercise/rehab as reported by each included paper. Understanding the unique strengths of AR games is critical to ensuring designers and therapists are equipped to maximize them in the context of ARREGs. Thus, future developers should seek to leverage and maximize these strengths for future ARREGs.

For AR:

- Keeps patient in contact with the real world, reducing cybersickness and minimizing interference with natural movement [4], [5], [24], [45], [47], [48]
- Real-time visual feedback of their real body, allowing users to self-correct [16], [20], [45], [46]
- Contactless tracking of hand and upper body resulting in more naturalistic movement and unencumbered proprioception [7], [23], [47], [48]
- Virtual obstacles are easier and safer to work with [54], [55]
- Low cost and convenient [20], [25], [45], [50]
- Variety of interaction methods [24]
- User interest and motivation increased due to novelty [24]
- Everybody can participate as a viewer (SAR) [10], or they can see each other as they play [53]
- Facilitates full-body interactions [11], [22]
- Can be done anywhere [52]
- Therapist presence is not necessary [49]
- Can record movement data for therapists [23], [24], [47]
- Therapists can provide remote exercises [50]

For games:

- Increased motivation & engagement despite repetitive tasks yielding greater compliance [10], [16], [20], [23]–[25], [29], [43], [44], [52], [55]
- Promote relatedness and socialization [10], [11]
- More authentic evaluations by pushing participants to try harder, reduce boredom, and distract from the evaluation itself [23], [50], [55]
- Can interweave learning and exercise [4], [16]

Combined, we see unique opportunities to provide rehabilitative or exercise experiences to more people (greater accessibility due to the low cost and high convenience of mobile AR), motivate patients to push themselves in both evaluation and therapy, a capacity to create unique social experiences, and several kinds of feedback be it via therapist or the game itself.

In contrast, very few limitations were expressed by the included works. General limitations surrounding AR are not unexpected and focus primarily on difficulty tracking complex user movements [46] (due to the nature of the exercise or ailment) or usability concerns. Designers should be mindful of their specific clinical population and ensure their ARREGs are tested with representative users.

3.3 User Representativeness & Diversity

Given that ARREGs are often made for specific, vulnerable groups (stroke, Parkinson's, etc.), we thought it was essential to ascertain whether the literature body has been successful in achieving studies with representative users, particularly as it has been shown that some disabilities also yield different experiences of presence in VR experiences [61]. Most VR research is done with the M-WEIRD population: Male, White, Educated, Industrialized, Rich, and Democratic [62]. This is disconcerting as the M-WEIRD population only accounts for 3% of the world's population, with the WEIRD population only accounting for 6% [63]. This would not be as concerning if these participants were not then touted as universal truths for everyone else [62]. Physical ability (or disability) is not an outlined population of the M-WEIRD group. However, we hypothesize that adding 'able-bodied' to the M-WEIRD considerations would further reduce how representative M-WEIRD is. We posit this because studies rarely include [31] or deliberately exclude participants with disabilities outside the context of health at the outset. We further posit that this phenomenon could feasibly be in the AR space, hence our inclusion of diversity concerns in this review.

This review found that most ARREGs focus on the rehabilitation side of health-related games, meaning that the works should inherently feature more differently-abled participants (and be more diverse) than non-health-related studies. I.e., to improve the scalability of findings, ARREGs should ideally be tested with the target demographic or the intended end user.

3.3.1 Gender, Age, Ethnicity, & Regions

Country	Paper #	Participant #	Total %
China	6	128	23.5
New Zealand	4	110	20.2
Australia	3	20	3.7
The Netherlands	3	60	11.0
Brazil	2	25	4.6
Canada	1	6	1.1
Greece	1	33	6.1
Poland	1	8	1.5
Portugal	1	18	3.3
Romania	1	84	15.4
South Korea	1	40	7.4
The United States	1	12	2.2

Table 1. Participants by country

Despite calls for greater demographic capturing of participants to lend more transparency to diversity (or lack thereof) in science [64], gender and age were the only reported demographic, with six studies not specifying any demographics (totaling 44 participants, ~8.1%). In line with a review of participant gender across VR studies [64], this work also found that < 1/4 of participants are female (118 or ~21.7%), with 240 (~44.1%) being male and 186 (~34.2%) unaccounted for.

15 studies reported age demographics (representing 491 participants or 90.3%), typically as ranges or averages. The youngest reported age is 6.5, with the oldest being 91. Six studies reported no demographics whatsoever, representing 44 participants (~8.1%).

No studies reported ethnicity. However, across all 25 papers, 12 countries are represented via the first authors (see Table 1), which denotes a more diverse participant pool than M-WEIRD.

We hypothesize that this participant pool is more diverse than those of non-medical contexts. However, there is still work concerning gender parity, greater demographic capturing, and transparency of demographics. Further, while data from 12 countries is a good start, more international diversity should be achieved, mainly as there is no way to ascertain participants' detailed ethnicity or nationality within each study.

3.3.2 Representative Clinical Participants

The target demographic refers to the clinical population to which the paper intends to provide its ARREG. The realized patient demographics denote whether their participant pool included the target population, with sixteen works able to sample participants from their intended use-case demographic (both clinical and non-clinical). Six works did not sample participants from their intended use case, with two not specifying whether their participants met clinical criteria. One work utilized usability experts and relevant therapists as a first step toward a clinical population.

Across all included works, there were 544 participants. 65 (~11.9%) were deliberate non-clinical controls, 127 (~23.3%) were not the target population and not deliberate controls, 310 were the target population (~56.9%) comprising 117 (~21.5%) general use participants, and 221 (~40.1%) clinical participants. General use denotes ARREGs that had no specified clinical group. 6 participants were usability experts, 3 were therapy experts, and 46 participants (~8.5%) were not specified. The represented clinical populations are as follows: Stroke (86), Parkinson's Disease (51), non-clinically defined anxiety (28), elderly institutionalized (18), Duchenne Muscular Dystrophy (16), power wheelchair users (5), manual wheelchair users (4), amblyopia (4).

The most common target groups were sufferers of stroke (thirteen studies) and Parkinson's disease (six studies), with four studies considering them simultaneously. We hypothesize this is due to three reasons:

1. The high global incidence of stroke (15 million each year; 5 million permanently disabled [65]) and the high incidence of Parkinson's Disease (~1 million in North America alone [66]), with both having higher disability rates and better outcomes with rehab,
2. A potential positive feedback loop wherein once something has been identified as promising in academic research, more research is done in that arena,
3. And the similar perceived presentations of both diseases [67]; further evidenced by one work referring to treating anyone needing upper arm motor rehabilitation and testing both stroke and Parkinson's patients [47].

Only two works [7], [24] had a formalized measure to assess the stage of disease progression and presentation in patients with Parkinson's disease [55]. This is critical to note as there is a wide range of disease presentations within one clinical user class, let alone across classes (stroke to Parkinson's, and so on. Evaluations such as this help bring transparency to the study and replicability (an imperative of science to mediate the replication crisis [68]). Further, this gives better context to the results as a game successful for a stage II Hoehn and Yahr [69] patient may be useless for a stage IV patient. Generally, these games should be guided by a therapist or medical professional that can help inform these nuanced but necessary distinctions.

3.4 ARREG Elements

3.4.1 Game Elements

Game elements refer to components and strategies used to "gamify," or transform, the exercise or rehab experience. Understanding leveraged elements presently gives future designers an idea of existing design metaphors to reference. Their incident rates can be found in Figure 2.

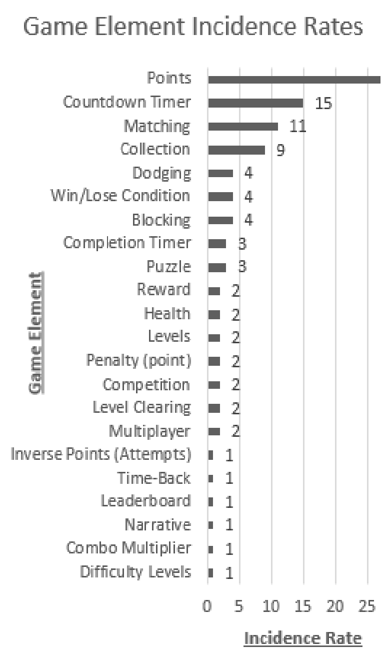

Figure 2. Game element incidences

Points refer to any form of scoring for performance. This includes successfully punching a moving object [4] or catching candy in a basket [23]. *Penalty Points* are point deductions for performance errors like failing to *block* an opponent's ball from scoring [16]. This also ties into *dodging,* where players must dodge incoming virtual artifacts like meteors [5]. *Health*, denoting some form of player health bar, was also used as a penalty in which the game ended once the bar hit 0 (this is also seen as a 'lose' condition). *Inverse Points* were used similarly to golf to capture attempts meaning that lower scores were better [51]. *Combo Multiplier* was used to reward players for continued good performance, meaning that the longer they did well, the higher their scores went up. *Leaderboards* were a formal roster of high scores accessible by the player; their high performance could put their name on the board [50]. Leaderboards can also be used for time-related mechanics.

Collection denotes any collection mechanic, like catching wine barrels in the river [10]. Collection was also used in conjunction with completion and countdown timers. *Completion Timers* are the time it takes to finish a task, whereas *Countdown Timers* denote that a player has a limited time to build their score as much as possible. *Time-Back* meant that players could add more time back to their countdown timers if they achieved an in-game goal, thus extending the game [22].

Matching denotes players needing to match two things together, like feeding the animal the correct food [20]. *Puzzles* denote any additional cognitive challenge, like solving math formulae, as a part of the game [25]. *Multi-player* means the game affords several players at once [53], with *competition* denoting players playing against each other. Narratives are a story-based rationale for playing the game, like good guys vs. bad guys [6]. *Win/Lose* indicators tell the player they succeeded (won) or failed (lost) at the game. *Rewards* are an additional means of motivating players for good game performance.

Levels mean that the game has several iterations of the task that can be played, whereas *Difficulty Levels* mean that players can select a difficulty level for their play experience. Of note, game element difficulty and extra element difficulty are different. The former refers to players being able to change the game's difficulty themselves. The latter refers to therapists being able to change the difficulty in the back end to fit the patient's needs more appropriately. Finally, *Level Clearing* means a player must reach a certain level of performance to make it to the next level (this could also be a 'win' condition).

By far, the most used game elements were points and timers. This makes sense as these are easy to implement in existing exercises. I.e., these can be just another means of motivation to complete as many reps as possible as quickly as possible. More subtle game elements that required more game-directed physical movement included collection, dodging, and blocking (which could also be used with points and timers). Surprisingly, matching was a highly prevalent form of game element, with many games requiring participants to play a memory game involving some reaching. Games using only a select few tended toward timers or points, essentially gamifying tasks instead of being serious games. More compelling game experiences could be fostered by utilizing more game elements than points and timers. Underexplored areas of interest are competition, leaderboards, and multi-player. These facets can facilitate more significant social interaction, positively affecting motivation and performance. However, these were among the least utilized game elements. Player autonomy in ARREGs could also be improved upon as there is more control given to the therapist in selecting difficulties than the patient. This could be helpful if patients are easily discouraged or prone to exercising dangerously. However, it is possible that players would be more motivated if given more control (perhaps on a case-by-case basis).

3.4.2 Rehab/Exercise Elements

Rehab/Exercise elements are the underlying rehab/exercise goals achieved through play. These focus on the physical activity being encouraged (e.g., reaching, pointing, grasping, stepping, etc.), the goal of the ARREG (to build strength, endurance, balance, and range), or an explanation of the targeted muscle groups (lower leg, forearm, upper limb, and the like). E.g., the player gets more points by gathering walnuts in a basket [24], and the process of catching walnuts is a non-trivial physical movement involving reach and speed. When designed thoughtfully, the games can elicit specific movements to mimic existing standard exercises. Alternatively, the games can simply get people moving in numerous ways to build strength and endurance.

Most papers focused on upper limb-related rehabilitation, with only six focusing on lower limbs or the entire body. This represents a dearth of research on lower-body-focused ARREGs. We hypothesize this could be due to three reasons:

1. The nature of the AR devices being utilized. Many are webcams situated on desktops with participant legs covered by the desktop,
2. AR devices are already built to track hands/arms,
3. Having arm-related abilities perhaps yields more practical independence,
4. The nature of the clinical population. Many sufferers of stroke and Parkinson's disease have problems with both their upper and lower limbs. This could make rehab with these clinical populations more dangerous (due to fall risk).

It was unanticipated how few works focused on exercise outside of rehabilitation. While exercise will not prevent Parkinson's disease, exercise is an essential protective agent for other forms of age-related physical disability. I.e., exergames could be used to prevent some types of injury that eventually require rehabilitation.

3.4.3 Extra Elements

Extra elements refer to any additional components added to enrich any ARREG experience to make them more effective, enjoyable, or better in some facet. These include auditory, visual, and haptic performance feedback, adaptability (or customizability), dual cognitive challenges, the ability to store data for therapists to view later, and more.

Some games featured simultaneous cognitive challenges, getting a 2-for-1 out of the patient's motivation and engagement. This is advantageous for clinical groups experiencing physical and cognitive disease pathologies like patients with stroke. Some systems were customizable by therapists, meaning that therapists could change the game's parameters to be more appropriate for a given patient. This drastically increases the range of possible players, making these games more employable. One game took it a step further and had an adaptive feature that automatically adjusted to the user after a short trial play, thus completely removing any therapist from the equation [11]. Finally, many ARREGs provided the ability to capture rehab-related data that could then be viewed by a therapist (and sometimes by the patient). These included capturing movement angles, velocity, time spent playing, accuracy, pinch width, and more. Finally, some extra elements included visual, auditory, and haptic elements to enhance the play experience further.

Extra elements bring ARREGs to the next level of employability and usefulness for patients and therapists. Extra elements, when possible, should be considered for future ARREG development.

3.5 Device and Modality

There was a wide range of devices from webcams and computers to mobile devices, Microsoft Kinects, projectors, HoloLenses, HTC Vive Pro's, projectors, and so on. There were also several combinations of different technologies, like webcams and Nintendo Wiimotes, AIRO II and mounted Leap Motion, etc. 15 works leverage video passthrough, five use optical see-through, and two use spatial augmented reality. Mobile AR devices were perceived as the lowest cost and most convenient.

AR has a greater range modality relative to VR. ARs particular unique strength is in spatial AR. Unfortunately, only two included works used spatial AR [10], [11]. Both SAR papers afforded the most social elements in their game experiences due to the AR being visible to anybody in the room. Generally, the spatial AR games were full-body exercises, whereas mobile phone games (except GeoBoids [22]) focused on the tiniest movements (reaching). Finally, webcam-based games featured predominantly medium movements (arm range of motion, reaching up high, etc.) and some small movements (hand reaching). Thus, perhaps future works aiming to work the entire body or provide a more significant social element should consider spatial AR modalities, while more minor to medium movements should consider video passthrough experiences. Moreover, spatial AR should be considered and evaluated more as it has a great, understudied capacity to afford some genuinely novel interactions and play experiences.

3.6 Measures and Outcomes

3.6.1 Rehab Measures

An essential facet in ARREG studies is measuring whether the ARREG has the intended rehabilitative effect. Twelve ARREGs used strictly physiological measures like muscle activation EMGs [16], movement speed [48], accuracy [46], distance [7], and tracking of patient movements with Kinect, Leap Motion, or other computer vision strategies to measure rehab effectiveness [5], [7], [23], [47]. Two works used therapist interpretation of movement to glean rehab effect [5], [29] or observations [11]. Only two works used a formal, validated set of examinations to measure physical gains (Fugl-Meyer Assessment [25] and Berg Balance Scale [55]). Most others used some degree of informed custom evaluation, like the duration of the session and the maximum reach distance paired with movement speed. Two works only used EMG data to determine whether the targeted muscle groups had been activated during play [20], [45].

Comparatively, nine works used strictly game-element-related performance measures for the same goal. These include points [6], [10], [22], [50], [52] completion time [18], [49], [51] health remaining at the end of the level [4], and the number of attempts to complete a level [18], [49].

Two works used both physiological and game-element-related measures to capture rehab effectiveness [24], [55]. The two remaining works used only participant perception of their experience but did not ask about their performance [53], [54].

Rehabilitative performance measures, regardless of form, were almost always happening simultaneously with gameplay. Thus, it is interesting that so few works included the game metrics of performance as part of their rehabilitative metrics of success. These metrics could be more understandable to a patient. A meaningful area of study then may be the further enmeshment of game elements with these more physical, formal performance measures. Any games using only game performance to measure rehabilitative performance should ensure that these game metrics are validated against intended outcomes. I.e., participants may be using improper form to achieve a higher point value, so it is important to have redundancy with other physical metrics of success.

3.6.2 Perception Measures

Given the importance of patient motivation in rehabilitative outcomes, capturing how user perceptions were being measured was meaningful. There was a mixed usage of either validated scales or custom questionnaires.

The most common perception measure came via some form of usability evaluation. Seven works used the *System Usability Scale* [70]. Another common validated scale, the Game Experience Questionnaire [71] was utilized four times. The final common validated perception measure was for cognitive load measured with the *NASA-TLX* [72], utilized by three works. Ten works created their own in-house questionnaires to mainly capture usability, followed by motivation, mental state, fun, and acceptance. Some studies were not specific in what exact questions were asked or the total amount of perceptions considered. Finally, three works completed semi-structured interviews or focus groups after the game experience. Some understanding of the assessed perceptions could only be gleaned via the reported outcomes of the work. Ultimately, these custom surveys were more likely to lack the necessary transparency for replication.

Other validated surveys included: 1) Adapted Presence Questionnaire, 2) World Health Organization Quality of Life, Satisfaction with Social Support Scale, 3) Basic Psychological Needs in Exercise Scale, Disabilities of Arm, Shoulder, and Hand (DASH) survey, 4) Technology Acceptance Questionnaire, 5) Barthdel Index, 6) State-Trait Anxiety Inventory (form Y), 7) Flow Short Scale, 8) Activities Specific Balance Confidence Scale, and 9) Simulator Sickness Questionnaire.

With this, we have a range from general usability studies (the majority) to some nuanced and detailed evaluations like the study that used the *World Health Organization Quality of Life, Satisfaction with Social Support scale*, and a *Basic Psychological Needs in Exercise scale* to determine how their game more deeply impacted their players [10]. While system usability is a critical facet in game development and poor usability is likely to impact motivation, it was interesting to see how few studies measured participant motivation despite claiming that improving motivation was the primary purpose of most ARREG papers. Further, it is unclear what value a usability evaluation has for a game designed for specific ailments (like Parkinson's) when the games are tested on non-clinical participants only.

Future works should endeavor to 1) use more formalized UX methods and ensure their participants are representative or well justified, 2) be transparent and share their custom usability questions to ensure greater comparison across studies, and 3) if a study claims they are interested in improving flow, motivation, and so on., then they should explore more than only usability, game experience, or cognitive load with participants. Just because a game is usable does not mean it is a motivating, trustworthy, flow-inducing, or successful at creating presence.

3.6.3 Outcomes

Outcomes were primarily positive, with some mixed results. Participants found the games interesting [20] and were motivated by them [20], [45], [46], [48], with ARREG users having significantly more motivation than control groups [29]. System usability was correlated with game experience [7]. Participants reported finding the games fun [11], [46], [49], fatiguing but not uncomfortable [20], and that the games provided a good workout [16]. Further, players report mostly good usability [6], [24] and experiencing flow [23]. Players found the games competitive but fair [11], with the games being easy to understand and learn [45], [24]. Finally, players reported significantly improved confidence [55] and less anxiety after ARREG intervention [4].

Some games had issues with their usability [23], [47], with others having some neutral scores [23] concerning engagement. In addition, two works report non-significant physical improvements pre/post-game intervention [18], [55], while another said it assessed motivation but did not report on motivation in the outcome [6].

Overall, ARREGs seem to be a great success from participants' perspectives. Interestingly, the literature seems to be split between validating that the games work (technically speaking), assessing their rehabilitative value, and determining player perceptions of the games. I.e., we found many works that evaluated pre/post performance on some physical dimension but did not evaluate user perceptions of the experience. We posit that these should go hand in hand. User perception alone is a powerful tool for understanding whether the ARREG will be used. However, if the game has poor to negligible rehabilitative gains, players could end up frustrated or spend time on the wrong thing. Worse, players could hurt themselves if the game challenges an unintended facet of the physical body. Conversely, knowing that the game has good to excellent rehabilitative results is compelling, but if players dislike the experience and are unwilling to stick with the game over time, then those results will never be realized.

4 Considerations For Future ARREGs

This review evaluated ARREGs that featured player perception evaluations. Overall, the games were well-received and seemingly effective at providing physical rehabilitation. This review provided some considerations for future ARREG development based on the current literature's successes, gaps, and weaknesses. Namely:

1. **Powering up Therapeutic Play Through Purposeful Development:** ARREGs should endeavor to understand the benefits of AR/games and leverage that as a grounding justification for using AR games (section 3.2) as their medium of intervention modality. These reasons should then inform the ARREG development. For a particularly compelling ARREG, perhaps the reasons for using AR and games overlap. Spatial AR seems uniquely situated to provide some genuinely novel play experiences and should be considered.
2. **Choose Your Character! Selecting the Appropriate AR Modality**: ARREGs should consider exploring more spatial AR opportunities, particularly for gross motor rehab/exercise experiences (section 3.5). Mobile AR should be considered when cost/convenience are a factor and fine motor movements are not the focus. Head-mounted displays work for medium and fine motor movement as they do not require the participant to hold a device. Further, most works focused on upper extremity rehab, making lower extremity rehab a ripe area for exploration.
3. **Diving into Diversity—The Power of Inclusion:** Diversity should be a primary consideration of researchers (section 3.3). At a minimum, researchers should report participants gender, age, and ethnicity (section 3.3.1). Ideally, researchers will also have a formalized means to measure player deficits and capacities to ensure consistency across groups of the same health class. In addition, researchers should strive to validate their games with the intended end user (section 3.3.2). Finally, researchers must ensure their participants are at parity with the group's gender norms. I.e., 50% unless the target clinical group has an established skew in gender prevalence. This would need to be justified in the paper. Therapists or medical professionals should be leveraged to guide meaningful ARREG development and testing.
4. **Assessing ARREGs for Gains AND Grins:** ARREGs should study the ARREG's ability to yield positive physical

gains *and* the user's experience of play (sections 3.6.1 and 3.6.2). Both components are a necessary consideration should the game hope to be implemented in therapy settings. There is no one-size-fits-all for humanity, and this is particularly true for the variable ways injuries and diseases can present. Further, game usability needs to be evaluated. Custom usability evaluations should be clearly explained with, ideally, the questions provided somewhere in the paper. More works should leverage existing, validated measures as outlined in section 3.6 or at least be guided by a subject-matter expert.

5. **Game On! Leveling Up Beyond Points and Timers:** ARREGs should endeavor to leverage more game elements than just points or timers as such experiences are mostly gamified exercises (section 3.4.1). Increased usage of game elements yielded more classically full, serious games. ARREGs can leverage the list of existing game elements to conceptualize future ARREGs, but should always consider new experiences. Further, social aspects appear to be an understudied area.
6. **Navigating New Realms of ARREG Potential:** Extra elements bring ARREG experiences to the next level of usefulness and immersion (section 3.4.3). ARREGs should endeavor to provide data to therapists and customization on the therapist's side to ensure games are appropriate for each user's abilities. Further, real-time feedback and audio make the play experience more meaningful and immersive.
7. **Gaming for Gains—Seamless Measurement:** Ideally, performance measures should happen simultaneously with gameplay to leverage the game's ability to push participants further [23] (section 3.6.1).
8. **Mind the Gap! Consider Underexplored ARREG Areas:** Exercise games seem understudied relative to rehabilitative games. The same is true for lower-body rehabilitation relative to upper-body rehabilitation and other clinical populations not yet represented (section 3.3).
9. **Ink the Details:** Finally, future published works should describe their games in enough detail to paint a picture of the game experience. Several works were eliminated from this review because no information was provided on the game experience.

To contextualize some of the aforementioned nine considerations, we point to an example paper: A particularly compelling game in relation to powering up therapeutic play through purposeful development, selecting and maximizing their AR modality, ensuring appropriate participant diversity, evaluating player perception using a customized series of questions that make sense to be customized based on their research questions, studying a unique series of questions and potential users, providing sufficient detail of their gameplay experience, and exemplifying some of the unique possibilities of ARREGs is that of IGYM [11]. This exercise game is provided via SAR on a gym floor. This game sought to bridge the inclusivity gap between differently abled players by creating a game everyone could play, seeking to facilitate both whole-body interactions and greater socialization. The game involved competition, points (via scoring), and blocking mechanics. The game was similar to *pong* in that players kicked a ball back and forth and tried to block it from scoring on their side. Players had an AR circle surrounding them that would 'bump' out or expand when the player made a kicking or arm-extending motion. The game offered adaptable puck speed, adaptable sizing of the circle, and a button to simulate the kick motion if needed. These adaptations could be changed between rounds and sets of players based on ability and desire.

5 LIMITATIONS & FUTURE DIRECTIONS

ARREGs are interdisciplinary, making it difficult to ascertain all relevant possible databases. This work is non-exhaustive; as such, other relevant ARREGS are likely not included in this evaluation. An expansion of this work to include more databases could be meaningful. Further, perhaps an extension of this work as a meta-analysis to truly understand the statistical findings of ARREGs is warranted. Strictly cognitive rehabilitative games were excluded from this work. However, they should also be evaluated in AR games as there is often an overlap between physical and cognitive rehabilitation (like in the context of stroke).

6 CONCLUSION

This work evaluated the existing ARREG literature to understand how these games are being furnished (AR modality and game elements), how the games are facilitating rehab/exercise (rehab/exercise elements), how performance is measured, user perceptions of the game experiences, who the games are being designed for and whether they are being tested with those groups, and the outcomes of existing ARREGs.

Broadly: The ARREG literature body features only 56.9% of the participants being appropriately representative users for the intended therapeutic populations, an underutilized variety of possible game elements, a lack of transparency in the descriptions of games in the literature, too many works using custom-made and non-validated usability measures despite there being several validated options, and an underwhelming amount of works that measure rehabilitative gains, usability, and patient perceptions all simultaneously. More positively, participants found the games engaging and motivating, this review found several interesting game elements for future consideration, and there was a range of clinical populations being developed for.

ARREGs are highly interdisciplinary and stand to provide profound gains in exercise and rehabilitation. To maximize future ARREG development, measurement, outcome reporting, literature cohesion, transparency, and participant diversity, this work offers some considerations for future developers. 1) Developers should maximize the unique benefits afforded by AR and games, 2) the AR modality should cater to the use-case (like SAR for gross motor movement and HMD for fine), 3) ARREGs should always be tested with the representative population and guided by a medical professional, 4) ARREGs should study the usability of the games, user perceptions of the experience, and the realized rehabilitative/exercise gains from using the game, 5) ARREGs should use more intensive game elements than simple points to afford more profound play experiences, 6) extra elements should be used where possible to deepen the game and therapeutic experience, 7) performance measures should happen simultaneous to gameplay where possible to maximize patient engagement with evaluation, 8) ARREGs should be mindful of gaps in the literature and seek to fill those gaps where possible, 9) future work should describe their games in sufficient detail to contextualize findings.

REFERENCES

[1] C. Stephanidis *et al.*, "Seven HCI Grand Challenges," *Int. J. Hum. Comput. Interact.*, vol. 35, no. 14, pp. 1229–1269, 2019, doi: 10.1080/10447318.2019.1619259.

[2] E. D. Mekler and K. Hornbæk, "Momentary pleasure or lasting meaning? Distinguishing eudaimonic and hedonic user experiences," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2016-Janua, pp. 4509–4520, 2016, doi: 10.1145/2858036.2858225.

[3] R.-G. Lee and S.-C. Tien, "Augmented Reality Game System Design for Stroke Rehabilitation Application," 2012, pp. 339–342, doi: 10.1109/CICSyN.2012.69.

[4] H. Yang, J. Li, Y. Bian, J. J. Liu, and J. J. Liu, "Development of An Indoor Exergame Based on Moving-target Hitting Task for COVID-19 Epidemic: A Comparison between AR and VR Modes," *IEEE Trans. Games*, vol. 14, no. 8, p. 1, 2021, doi: 10.1109/TG.2021.3118035.

[5] A. Nowak, M. Wozniak, M. Pieprzowski, and A. Romanowski, "Towards Amblyopia Therapy Using Mixed Reality Technology," in *2018 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 2018, pp. 279–282.

[6] S. L. Kim, H. J. Suk, J. H. Kang, J. M. Jung, T. H. Laine, and J. Westlin, "Using Unity 3D to facilitate mobile augmented reality game development," in *2014 IEEE World Forum on Internet of Things (WF-IoT)*, 2014, pp. 21–26, doi: 10.1109/WF-IoT.2014.6803110.

[7] M. A. Cidota, P. J. M. Bank, P. W. Ouwehand, and S. G. Lukosch, "Assessing Upper Extremity Motor Dysfunction Using an Augmented Reality Game," in *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2017, pp. 144–154, doi: 10.1109/ISMAR.2017.31.

[8] Y. A. Sekhavat and M. S. Namani, "Projection-Based AR: Effective Visual Feedback in Gait Rehabilitation," *IEEE Trans. Human-Machine Syst.*, vol. 48, no. 6, pp. 626–636, Dec. 2018, doi: 10.1109/THMS.2018.2860579.

[9] A. Voida and S. Greenberg, "Wii all play: The console game as a computational meeting place," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1559–1568, 2009, doi: 10.1145/1518701.1518940.

[10] H. Cardoso, A. Bernardino, M. Sanches, and L. L. Loureiro, "Exergames and their benefits in the perception of the Quality of Life and Socialization on institutionalized older adults," in *2019 5th Experiment International Conference (exp.at'19)*, 2019, pp. 298–304, doi: 10.1109/EXPAT.2019.8876469.

[11] R. Graf *et al.*, "IGYM: An Interactive Floor Projection System for Inclusive Exergame Environments," in *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 2019, pp. 31–43, doi: 10.1145/3311350.3347161.

[12] K. Harris and D. Reid, "The influence of virtual reality play on children's motivation Key," no. 1, pp. 21–29, 2005.

[13] A. M. M. Funabashi, R. V Aranha, T. D. Silva, C. B. M. Monteiro, W. S. Silva, and F. L. S. Nunes, "AGaR: A VR Serious Game to Support the Recovery of Post-Stroke Patients," in *2017 19th Symposium on Virtual and Augmented Reality (SVR)*, 2017, pp. 279–288, doi: 10.1109/SVR.2017.15.

[14] E. C. S. Murphy, L. Carson, W. Neal, C. Baylis, D. Donley, and R. Yeater, "Effects of an exercise intervention using Dance Dance Revolution on endothelial function and other risk factors in overweight children," *Int. J. Pediatr. Obes.*, vol. 4, no. 4, pp. 205–214, 2009, doi: 10.3109/17477160902846187.

[15] J. Sohnsmeyer, H. Gilbrich, and B. Weisser, "Effect of a six-week-intervention with an activity-promoting video game on isometric muscle strength in elderly subjects," *Int. J. Comput. Sci. Sport*, vol. 9, no. 2: Special Edition, pp. 75–80, 2009.

[16] A. Dinevan, Y. M. Aung, and A. Al-Jumaily, "Human computer interactive system for fast recovery based stroke rehabilitation," in *2011 11th International Conference on Hybrid Intelligent Systems (HIS)*, 2011, pp. 647–652, doi: 10.1109/HIS.2011.6122182.

[17] D. Michael and S. Chen, *Serious Games: Games that Educate, Train, and Inform*, 1st ed. Cengage Learning PTR; 1st edition, 2005.

[18] M. King, L. Hale, A. Pekkari, M. Persson, M. Gregorsson, and M. Nilsson, "An affordable, computerised, table-based exercise system for stroke survivors," *i-CREATe 2009 - Int. Conv. Rehabil. Eng. Assist. Technol.*, vol. 5, no. 4, pp. 288–293, Jul. 2010.

[19] J. Ku, Y. J. Kim, S. Cho, T. Lim, H. S. Lee, and Y. J. Kang, "Three-dimensional augmented reality system for balance and mobility rehabilitation in the elderly: A randomized controlled trial," *Cyberpsychol. Behav. Soc. Netw.*, vol. 22, no. 2, pp. 132–141, Feb. 2019, doi: 10.1089/cyber.2018.0261.

[20] Y. M. Aung and A. Al-Jumaily, "Rehabilitation exercise with real-time muscle simulation based EMG and AR," in *2011 11th International Conference on Hybrid Intelligent Systems (HIS)*, 2011, pp. 641–646, doi: 10.1109/HIS.2011.6122181.

[21] D. Zhang, Y. Shen, S. K. Ong, and A. Y. C. Nee, "An Affordable Augmented Reality Based Rehabilitation System for Hand Motions," in *2010 International Conference on Cyberworlds*, 2010, pp. 346–353, doi: 10.1109/CW.2010.31.

[22] R. W. Lindeman, G. Lee, L. Beattie, H. Gamper, R. Pathinarupothi, and A. Akhilesh, "GeoBoids: A mobile AR application for exergaming," in *2012 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities (ISMAR-AMH)*, 2012, pp. 93–94, doi: 10.1109/ISMAR-AMH.2012.6483997.

[23] M. A. Cidota, P. J. M. Bank, and S. G. Lukosch, "Design Recommendations for Augmented Reality Games for Objective Assessment of Upper Extremity Motor Dysfunction," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019, pp. 1430–1438, doi: 10.1109/VR.2019.8797729.

[24] P. J. M. M. Bank, M. A. Cidota, P. W. (Elma) W. Ouwehand, and S. G. Lukosch, "Patient-tailored augmented reality games for assessing upper extremity motor impairments in Parkinson's disease and stroke," *J. Med. Syst.*, vol. 42, no. 12, pp. 246 (11 pp.)-, 2018, doi: 10.1007/s10916-018-1100-9.

[25] C. Li *et al.*, "Long-term effectiveness and adoption of a cellphone augmented reality system on patients with stroke: Randomized controlled trial," *JMIR Serious Games*, vol. 9, no. 4, p. e30184, Nov. 2021.

[26] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, "Augmented reality: A class of displays on the reality-virtuality continuum," *Telemanipulator Telepresence Technol.*, vol. 2351, no. January, pp. 282–292, 1995, doi: 10.1117/12.197321.

[27] R. T. Azuma, "A survey of augmented reality," *Found. Trends Human-Computer Interact.*, vol. 6, no. 4, pp. 335–385, 1997, doi: 10.1561/1100000049.

[28] N. Faric *et al.*, "What players of virtual reality exercise games want: Thematic analysis of web-based reviews," *J. Med. Internet Res.*, vol. 21, no. 9, pp. 1–13, 2019, doi: 10.2196/13833.

[29] A. G. Dionísio Corrêa, I. K. Ficheman, R. De Deus Lopes, A. N. Klein, and S. J. Nakazune, "Augmented Reality in Occupacional Therapy," in *2013 8th Iberian Conference on Information Systems and Technologies (CISTI)*, 2013, pp. 1–6.

[30] S. N. Junaini, A. A. Kamal, A. H. Hashim, N. M. Shaipullah, and L. Truna, "Augmented and Virtual Reality Games for Occupational Safety and Health Training: A Systematic Review and Prospects for the Post-Pandemic Era," *Int. J. online Biomed. Eng.*, vol. 18, no. 10, pp. 43–63, 2022, doi: 10.3991/ijoe.v18i10.30879.

[31] O. Mubin, F. Alnajjar, N. Jishtu, B. Alsinglawi, and A. Al Mahmud, "Exoskeletons with virtual reality, augmented reality, and gamification for stroke patients' rehabilitation: Systematic review," *JMIR Rehabil. Assist. Technol.*, vol. 6, no. 2, pp. 1–11, 2019, doi: 10.2196/12010.

[32] V. C. Cavalcanti, M. I. De Santana, A. E. F. Da Gama, and W. F. M. Correia, "Usability assessments for augmented reality motor rehabilitation solutions: A systematic review," *Int. J. Comput. Games Technol.*, vol. 2018, 2018.

[33] J. Yim and T. C. N. Graham, "Using games to increase exercise motivation," *Proc. 2007 Conf. Futur. Play. Futur. Play '07*, pp. 166–173, 2007, doi: 10.1145/1328202.1328232.

[34] M. J. V. Gil *et al.*, "Augmented reality in physical therapy: Systematic review and meta-analysis," *JMIR Serious Games*, vol. 9, no. 4, p. e30985, Dec. 2021, doi: 10.2196/30985.

[35] F. U. Yu, H. U. Yan, and V. Sundstedt, "A Systematic Literature Review of Virtual, Augmented, and Mixed Reality Game Applications in Healthcare," *ACM Trans. Comput. Healthc.*, vol. 3, no. 2, 2022, doi: 10.1145/3472303.

[36] A. Liberati *et al.*, *The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration*, vol. 62, no. 10. 2009.

[37] O. Mubin, F. Alnajjar, N. Jishtu, B. Alsinglawi, and A. Al Mahmud, "Exoskeletons with virtual reality, augmented reality, and gamification for stroke patients' rehabilitation: Systematic review," *JMIR Rehabil. Assist. Technol.*, vol. 6, no. 2, p. e12010, Sep. 2019, doi: 10.2196/12010.

[38] A. de Miguel-Rubio, M. Dolores Rubio, A. Alba-Rueda, A. Salazar, J. A. Moral-Munoz, and D. Lucena-Anton, "Virtual reality systems for upper limb motor function recovery in patients with spinal cord injury: Systematic review and meta-analysis," *JMIR mHealth uHealth*, vol. 8, no. 12, pp. 1–15, 2020, doi: 10.2196/22537.

[39] G. L. Dimaguila, K. Gray, and M. Merolli, "Person-generated health data in simulated rehabilitation using kinect for stroke: Literature review," *JMIR Rehabil. Assist. Technol.*, vol. 20, no. 5, pp. 1–12, 2018, doi: 10.2196/rehab.9123.

[40] B. Steiner, L. Elgert, B. Saalfeld, and K. H. Wolf, "Gamification in rehabilitation of patients with musculoskeletal diseases of the shoulder: Scoping review," *JMIR Serious Games*, vol. 8, no. 3, 2020, doi: 10.2196/19914.

[41] D. Johnson, S. Deterding, K. A. Kuhn, A. Staneva, S. Stoyanov, and L. Hides, "Gamification for health and wellbeing: A systematic review of the literature," *Internet Interv.*, vol. 6, pp. 89–106, 2016, doi: 10.1016/j.invent.2016.10.002.

[42] D. Moher *et al.*, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *PLoS Med.*, vol. 6, no. 7, 2009, doi: 10.1371/journal.pmed.1000097.

[43] E. Elsamahy, "an Investigation Into Using Digital Games-Based Learning in Architecture Education," *Archit. Plan. J. - Spec. Ed.*, vol. SPECIAL ED, no. March, 2017.

[44] S. Deterding, M. Sicart, L. Nacke, and ... K. O., "Gamification: Using Game Design Elements in Non-Gaming Contexts," in *CHI'11*, 2011, pp. 5–8.

[45] Y. M. Aung and A. Al-Jumaily, "AR based upper limb rehabilitation system," in *2012 4th IEEE RAS EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2012, pp. 213–218, doi: 10.1109/BioRob.2012.6290680.

[46] R. R. Barioni, T. M. Chaves, L. Figueiredo, V. Teichrieb, E. V. Neto, and A. E. F. Da Gama, "ARkanoidAR: An Augmented Reality System to Guide Biomechanical Movements at Sagittal Plane," in *2017 19th Symposium on Virtual and Augmented Reality (SVR)*, 2017, pp. 207–214, doi: 10.1109/SVR.2017.34.

[47] J. Goderie *et al.*, "[POSTER] ChiroChroma: An Augmented Reality Game for the Assessment of Hand Motor Functionality," in *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, 2017, pp. 115–120, doi: 10.1109/ISMAR-Adjunct.2017.44.

[48] M. Hoda, H. Dong, D. Ahmed, and A. El Saddik, "Cloud-based rehabilitation exergames system," in *2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, 2014, pp. 1–6, doi: 10.1109/ICMEW.2014.6890660.

[49] S. Hoermann, J. Collins, and H. Regenbrecht, "Learning from rehabilitation: A bi-manual interface approach," in *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, 2015, pp. 163–164, doi: 10.1109/3DUI.2015.7131751.

[50] D. Koulouris and A. Menychtas, "An IoT-Enabled Platform for the Assessment of Physical and Mental Activities Utilizing Augmented Reality Exergaming," pp. 1–19, 2022.

[51] H. Regenbrecht *et al.*, "Out of reach? — A novel AR interface approach for motor rehabilitation," in *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, 2011, pp. 219–228, doi: 10.1109/ISMAR.2011.6092389.

[52] X. Song, L. Ding, J. Zhao, J. Jia, and P. Shull, "Cellphone Augmented Reality Game-based Rehabilitation for Improving Motor Function and Mental State after Stroke," in *2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2019, pp. 1–4, doi: 10.1109/BSN.2019.8771093.

[53] T. Y. Tang, P. Winoto, and R. Y. Wang, "Supporting Multiplayer Online Ball Passing at a Distance with Multiple Sets of Kinect," *2015 Second Int. Conf. Soft Comput. Mach. Intell.*, pp. 107–111, 2015, doi: 10.1109/ISCMI.2015.40.

[54] H. Tsai and C. Chen, "Usability Evaluation of a Muscle Strength Training System for Older Adults," *2022 Int. Conf. Comput. Appl.*, pp. 1–5, doi: 10.1109/ICCA56443.2022.10039541.

[55] Y.-W. Wang, C.-H. Chen, and Y.-C. Lin, "Balance Rehabilitation System for Parkinson's Disease Patients based on Augmented Reality," in *2020 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE)*, 2020, pp. 191–194, doi: 10.1109/ECICE50847.2020.9302018.

[56] O. Marquet, C. Alberico, D. Adlakha, and J. A. Hipp, "Examining motivations to play pokémon go and their influence on perceived outcomes and physical activity," *JMIR Serious Games*, vol. 5, no. 4, 2017, doi: 10.2196/games.8048.

[57] P. Rasche, A. Schlomann, and A. Mertens, "Who is still playing pokémon go? a web-based survey," *JMIR Serious Games*, vol. 5, no. 2, pp. 1–14, 2017, doi: 10.2196/games.7197.

[58] T. Althoff, R. W. White, and E. Horvitz, "Influence of pokémon go on physical activity: Study and implications," *J. Med. Internet Res.*, vol. 18, no. 12, 2016, doi: 10.2196/jmir.6759.

[59] J. Arjoranta, T. Kari, and M. Salo, "Exploring features of the pervasive game pokémon GO that enable behavior change: Qualitative study," *JMIR Serious Games*, vol. 8, no. 2, pp. 1–13, 2020, doi: 10.2196/15967.

[60] A. K. Lindqvist, D. Castelli, J. Hallberg, and S. Rutberg, "The praise and price of pokémon go:a qualitative study of children's and parents' experiences," *JMIR Serious Games*, vol. 20, no. 1, pp. 1–7, 2018, doi: 10.2196/games.8979.

[61] R. Guo, G. Samaraweera, and J. Quarles, "The Effects of VEs on Mobility Impaired Users: Presence, Gait, and Physiological Response," in *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*, 2013, pp. 59–68, doi: 10.1145/2503713.2503719.

[62] J. Henrich, S. J. Heine, and A. Norenzayan, "Weird people," *Behav. Brain Sci.*, vol. 33, no. 2–3, pp. 61–135, 2010.

[63] T. C. Peck, K. A. McMullen, J. Quarles, K. Johnsen, C. Sandor, and M. Billinghurst, "DiVRsify: Break the Cycle and Develop VR for Everyone," *IEEE Comput. Graph. Appl.*, vol. 41, no. 6, pp. 133–142, 2021, doi: 10.1109/MCG.2021.3113455.

[64] T. C. Peck, L. E. Sockol, and S. M. Hancock, "Mind the Gap: The Underrepresentation of Female Participants and Authors in Virtual Reality Research," *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 5, pp. 1945–1954, 2020, doi: 10.1109/TVCG.2020.2973498.

[65] World Health Organization (WHO), "Stroke , Cerebrovascular accident." .

[66] C. Marras *et al.*, "Prevalence of Parkinson's disease across North America," *npj Park. Dis.*, vol. 4, no. 1, pp. 1–7, 2018, doi: 10.1038/s41531-018-0058-0.

[67] R. Caslake, K. S. M. Taylor, and C. E. Counsell, "Parkinson's disease misdiagnosed as stroke," *BMJ Case Rep.*, pp. 7–9, 2009, doi: 10.1136/bcr.07.2008.0558.

[68] H. Pashler and C. R. Harris, "Is the Replicability Crisis Overblown? Three Arguments Examined," *Perspect. Psychol. Sci.*, vol. 7, no. 6, pp. 531–536, 2012, doi: 10.1177/1745691612463401.

[69] M. M. Hoehn and M. D. Yahr, "Parkinsonism : onset , progression , and mortality," vol. 17, no. May, 1967.

[70] J. Brooke, "SUS: A 'Quick and Dirty' Usability Scale," *Usability Eval. Ind.*, no. January 1996, pp. 207–212, 2020, doi: 10.1201/9781498710411-35.

[71] K. Poels, Y. A. W. de Kort, and W. A. IJsselsteijn, "Game Experience Questionnaire The fun of gaming: Measuring the human experience of media enjoyment," no. 2007, 2007.

[72] S. G. Hart and L. E. Staveland, "Development of Nasa Tlx (Task Load Index): Results of Empirical and Theoretical Research. Human Mental Workload," *Hum. Ment. Workload*, pp. 139–183, 1988.