Erschienen in: Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization. - New York, NY, USA : ACM, 2023. - S. 283-290. - ISBN 978-1-4503-9891-6
https://doi.org/10.1145/3563359.3597389

Group Adapted Avatar Recommendations for Exergames

Robin Cromjongh
Utrecht University
Utrecht, The Netherlands

Quentin van Reenen
Utrecht University
Utrecht, The Netherlands

Laura König
University of Bayreuth
Bayreuth, Germany

Martina Kanning
University of Konstanz
Konstanz, Germany

Ulf-Dietrich Reips
University of Konstanz
Konstanz, Germany

Tiare Feuchtner
University of Konstanz
Konstanz, Germany

Hanna Hauptmann
Utrecht University
Utrecht, The Netherlands

Figure 1: Posed scenarios of two people playing VR exergames. Left: The display in the back shows a spectator's perspective of an exemplary VR Squash game (captured from [11]), illustrating how avatar appearance might differ to reflect the players' skill, fitness level, or in-game abilities. Right: The inset picture shows a player's first-person perspective of their opponent's avatar in a virtual Ping Pong simulation (captured from [14]).

ABSTRACT

Exergames are a promising way to encourage physical activity in the population. Especially competitive gaming has been shown to boost physical activity during gameplay. However, differences in physical abilities and fitness can lead to anxiety, fear of failure, or frustration. One way to mitigate these inhibitors is to balance the exergaming difficulty between competing players. This paper investigates the expectations and attitudes towards adaptivity in sports games, both in real life and with digital support. To that end, we present a survey with 421 participants investigating the general reaction to group adaptivity in sports games as well as a focus group discussing the reactions to group adaptive avatar recommendations within the game Mario Tennis Aces. Our results show that there is potential for group adaptive exergames to increase engagement, especially for non-sporty and female users, and that the first prototypical implementation was perceived positively regarding fairness and expected physical activity.

KEYWORDS

exergame, group recommendation, personalization, physical activity, avatar, fairness

1 INTRODUCTION

The frequent lack of physical activity in the population can lead to a higher risk of non-communicable diseases and increased healthcare costs. Serious games, gamification, persuasive applications, and behaviour change applications have all shown promise to increase physical activity. One specific form of gameful physical activity, so-called “exergames”, can make physical exertion more attractive through video (game) simulations or provide additional health-promoting measures in existing popular games. While exergames originated on entertainment platforms that were played on the living room TV with a controller, like the Wii or Playstation Move, recent technological advances have also made virtual reality (VR) headsets available and affordable. This offers new opportunities for exergames in terms of immersion and user tracking (see exemplary illustration in Figure 1), which allows closer analysis of user performance and facilitates multimodal manipulation of the users’ perceptions that affect their overall experience. As exergames often involve a strong competitive aspect (e.g., beating another player or a previous high score), this research focuses on multi-player interaction and, in particular, on group practices in exergames. The central research questions are (1) how the exergame can adapt to a player’s physical characteristics, skill level, and fitness in a group context, and (2) how this may impact the overall gaming experience. For example, a less fit, physically smaller, or less experienced player may be supported by the system to compete more fairly with a player who has physical advantages or greater skill. We hypothesize that adjusting the challenge to each player’s personal level of ability and fitness has the potential to increase motivation, strengthen the experience of flow and improve the overall training outcome. In this paper, we want to investigate the practices and issues in real-life group sports activities and the potential of digital support in adapting to a group of differently skilled co-players. We present a survey with 421 participants on their group sports attitudes and their perception of digital adaptivity support. We further built a prototypical group recommender system for the commercial exergame Mario Tennis Aces and tested the perception of this recommendation with a focus group. In summary, our contributions are 1) the transfer of real-life group sports attitudes and practices to exergames, 2) identifying potential target groups for digital group adaptivity support, and 3) testing perception of and attitude towards providing digital group adaptivity support in the form of avatar recommendations for exergames.

2 RELATED WORK

Digital interventions to promote physical activity have the potential to effectively promote healthier lifestyles [7, 24]. However, these interventions are currently used by a relatively small proportion of the population, and many users disengage after a short amount of time, which is likely insufficient to achieve long-lasting behaviour change [2, 13]. Exergames are one such intervention that has been shown to have positive effects on fitness, muscle strength [23], and also motivation [17]. In particular, the increasingly sophisticated technical possibilities in virtual reality offer an effective platform for exergames [26]. It is already known from other serious games that personalisation can increase the effectiveness of the game compared to “one-size-fits-all” approaches. Studies show that overweight players, in particular, have a different gaming experience and effects from playing exergames than normal-weight players [1]. In the context of these differences, compensation mechanisms based on actual effort during gameplay have already been investigated [4]. It should be noted that actual and perceived effort may differ depending on the game [27]. To some extent, such differences in perception can be used to obtain comparable fitness improvements under lower perceived effort. In VR, players are often represented as avatars, which we can embody under certain conditions (“embodiment”) and even accept the artificial/virtual body as our own to some degree (“body ownership”) [12]. This embodiment influences our self-perception (“body image” and “body schema” [9]), perception of the world, and, as a result, our behaviour. In terms of body size and weight, for example, it has been shown that viewing an overweight or underweight virtual avatar from a first-person perspective affects how players perceive their own bodies [18]. Such effects can also be achieved through the non-visual representation of the players, such as the sound of their own footsteps [22]. In general, playing in teams is conducive to exergame effects [8]. In particular, the superiority and similarity of the partner play a role in increasing motivation [8]. Thus, with respect to the adaptive representation of the avatar, the mutually perceived competitiveness should also be considered. A solution to the dilemma of fairness and challenge is provided by the sequential switching of recommended avatars within the group. Here, the first step is to combine the individual evaluations of the available avatars into a group recommendation according to different aggregation strategies [15]. The resulting suboptimal aggregations can be balanced in the sequence of game rounds by iteratively re-evaluating the needs of players based on the preceding game experience, the memorability of this experience, and the emotional contagion from other players [16].

3 SURVEY ON GROUP ADAPTIVE EXERGAMES

This survey was pre-registered on the Open Science Framework (OSF) prior to data collection to enhance transparency and rigour [3]. Within the survey, this paper focuses on the following research questions: “Is adaptation in group sports activities needed and expected?” and “Is the adaptation of exergames to individual skills perceived to be motivating and fair?” The broader aim of the survey is to understand the (long-term) use and impact of different fitness apps, fitness trackers, and nutrition apps, as well as their dependency on demographics, personality traits, motivational attitudes, and exercise patterns. The overall survey aims to improve digital interventions for promoting healthier lifestyles.

3.1 Study Procedure

We conducted an online survey from October 2022 to January 2023 using Tivian Unipark[1] Participation in the survey was expected to take around 15 minutes, including reading the study description, completing the written informed consent, answering the questionnaire, and a short debriefing at the end of the survey. The first block of questions asked for sociodemographic information as well as the Big 5 and social comparison orientation personality traits

[1]Tivian Unipark: https://www.unipark.com/en/

[10, 19, 20]. The second block focused on the ownership and use of fitness apps, fitness trackers, and nutrition apps [13], as well as the participant's goals in using them. The third block assesses the physical activity patterns of the participant and their attitude towards adaptivity in group sports activities and exergames. In this paper, we especially focus on the latter section of the survey regarding adaptive sporting and exergaming, which asks for the participant's opinions on two randomized group sporting scenarios and on the perceived effect of digital group adaptivity support using 7-point Likert scales. We will further include gender as a relevant dimension from the sociodemographic data and the presence of regular sporting behaviour from the data on physical activity patterns. More specifically, we asked our questions as follows (scenarios A and B were randomized in order):

Gender and regular sporting behaviour:

(1) You are: female, male, diverse
(2) Do you regularly play individual sports (e.g. athletics, swimming) or team sports (e.g. football, volleyball)? This includes competitive sports (e.g. competitions, playing in a team) as well as recreational sports (e.g. playing Frisbee with friends): yes, no

A: Imagine that a friend asks you to play sports together (e.g. ball sports like volleyball or tennis). You are aware that your friend and possibly the other players are better at the sport than you are.

(1) How likely is it that you will accept the invitation?: Very unlikely ... Very likely
(2) If I accept the challenge, I expect the players who are superior to me to adapt to my abilities: Not agree at all ... Fully Agree

B: Imagine that a friend asks you to play a sport together (e.g., ball sports like volleyball or tennis). You are aware that you are much better at the sport than your friend and possibly the other players.

(1) How likely is it that you will accept the invitation?: Very unlikely ... Very likely
(2) If I accept the prompt, I will adapt my playing style to the skills of my teammates: Not agree at all ... Fully Agree

Now imagine that there is a system that can match the playing abilities of the users. This could be, for example, an adaptation of the rules of the game (weaker players get more attempts at mini-golf) or a digital adaptation (virtual reality tennis with different racket sizes).

(1) An adaptation to my personal abilities would motivate me to do more sport: Not agree at all ... Fully Agree
(2) An adaptation to my personal abilities would be perceived as fair for all players: Not agree at all ... Fully Agree

3.2 Participants

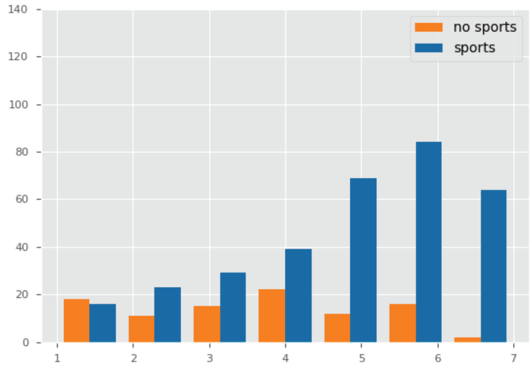
The selection criteria included participants with a reported age above 18, sufficient German language skills, and ownership of a mobile or wearable device. Recruitment was conducted via social media and local recruitment at different German universities. The aim of 400 participants should allow detection of small to medium effect sizes (e.g., Pearson's r = 0.14, Cohen's d = 0.28; $\alpha = 0.05$; [5]) at 80% power. The final participant sample consisted of 421 answers (341 reportedly female) with an average reported age of 27,5 (min 18, max 68). Due to the recruitment procedure targeting people using lifestyle apps and/or fitness trackers and the general bias of participants in health-related surveys, our samples show a clear bias towards female participants (341/421) and participants who regularly do sports (324/421).

3.3 Results

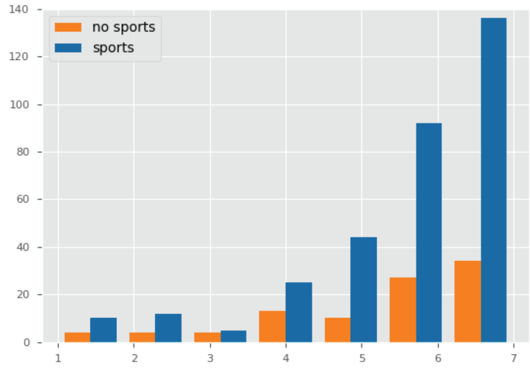
None of the resulting data from the different 7-point Likert scale answers was normally distributed. We thus chose to analyse the data with Wilcoxon signed-rank tests (for matched pairs) and Wilcoxon rank-sum tests (for unmatched groups). In the following, we first present results regarding the differences between willingness to join group sports as the least skilled player or the most skilled player and its dependency on regular sporting behaviour (Figure 2). Second, we look at how the perceived motivational boost and the perceived fairness of digital adaptivity to skill levels in group exergaming depend on regular sporting behaviour, motivation to join as the least skilled player, and gender (Figure 3).

Figure 2 reveals clear differences in both the willingness to join a group sports game and the expectation that the superior players would adapt their behaviour to co-players between the scenario where the participant is the least skilled player or the most skilled player. Participants were significantly more willing to join a group sports activity as the most skilled player than as the least skilled player (Wilcoxon signed-rank test, T=3992.5, $p<0.001$). The same difference holds true for both participants who regularly do sports (Wilcoxon signed-rank test, T=2884.0, $p<0.001$) and participants who don't regularly do sports (Wilcoxon signed-rank test, T=77.0, $p<0.001$). At the same time, the willingness to join as the least skilled player was significantly lower for participants who don't regularly do sports than for participants who regularly do sports (Wilcoxon rank-sum test, Z=-6.326, $p<0.001$). This impact of sport had no significant effect on the willingness to join as the most skilled player (Wilcoxon rank-sum test, Z=-1.511, p=0.131). When looking at the expectations regarding adaptive behaviour in group sports activities, we see that, while people mainly would be willing to adapt their own behaviour towards less skilled co-players, they do not expect others to do the same for them: Participants were significantly more willing to adapt their own behaviour in a group sports activity as the most skilled player than they expected others to adapt their behaviour when they were the least skilled player (Wilcoxon signed-rank test, T=3354.0, $p<0.001$). This difference holds true both for participants who regularly do sports (Wilcoxon signed-rank test, T=2099.0, $p<0.001$) and participants who don't regularly do sports (Wilcoxon signed-rank test, T=147.5, $p<0.001$). The expectation of adaptation is not significantly impacted by regular sporting behaviour in either the least skilled player scenario (Wilcoxon rank-sum test, Z=-0.569, p=0.569) or the most skilled player scenario (Wilcoxon rank-sum test, Z=-0.357, p=0.721)

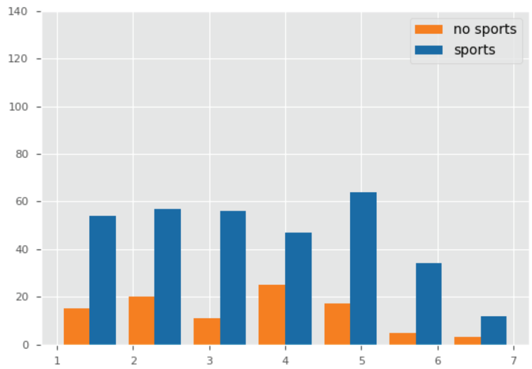
As shown in Figure 3, digital support in adapting to skills within a group received mixed ratings for both the perceived motivational boost and the perceived fairness. As a relevant factor from real-world behaviour, we first looked into differences between people who regularly do sports and those who do not. However, the regularity of doing sports had no significant impact on perceived



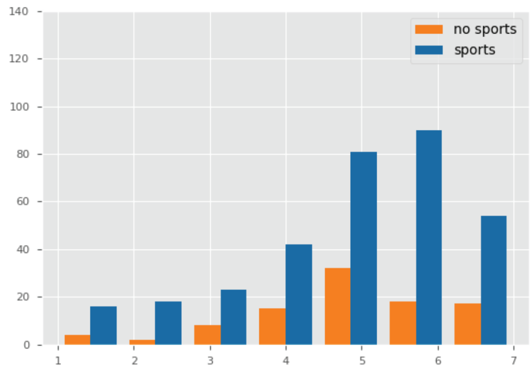
(a) Willingness to join as least skilled player



(b) Willingness to join as most skilled player



(c) Expectation of others adapting to least skilled player



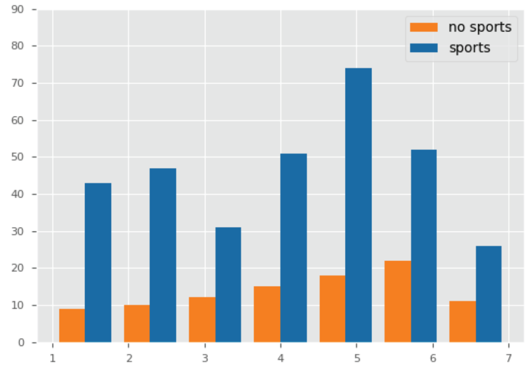
(d) Expectation of adapting to others as most skilled player

Figure 2: These graphs show the differences between the willingness to join a group sports activity (top) as the least skilled player (left) or the most skilled player (right) as well as the participants' expectation of the more skilled players adapting their behaviour to the others (bottom) and the dependency in each of these cases on the participant having a regular sporting behaviour (blue) or not (orange).

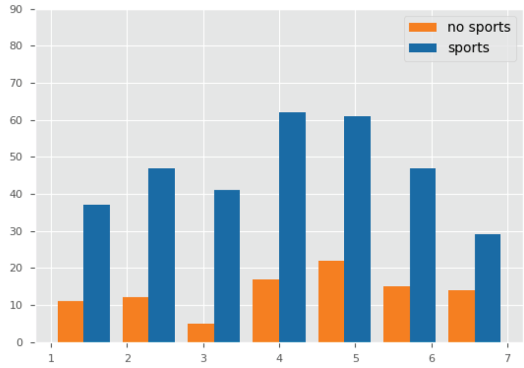
motivational boost (Wilcoxon rank-sum test, Z=1.688, p=0.091) or perceived fairness (Wilcoxon rank-sum test, Z=1.668, p=0.095). To go one step further into especially the group that needs additional motivation to join group sports activities, we next split the sample into participants willing to join even as the least skilled player (Rating >= 4) and participants not willing to join such activities. For this distinction, we see a significantly higher perceived motivational boost from digital adaptivity support for participants not willing to join, which is exactly the group we aim to motivate (Wilcoxon rank-sum test, Z=3.077, p=0.002). There was no comparable significant impact on the perceived fairness (Wilcoxon rank-sum test, Z=1.784, p=0.074). Finally, we looked into gender differences in perceiving digital adaptivity support. Women report a significantly higher perceived motivational boost with digital adaptivity support (Wilcoxon rank-sum test, Z=-4.484, p<0.001) and significantly higher perceived fairness in such a system (Wilcoxon rank-sum test, Z=-2.163, p=0.031) than men.

Based on the potential for adaptive group exergames shown in the survey results, we want to gain insights into the perception of real users when presenting them with gaming avatars adapted to their physical activity level compared to their co-players. To this end, we conducted a focus group in the Netherlands in March 2023. The game used in this research is Mario Tennis Aces[2], a tennis-like game created for the Nintendo Switch game console. Instead of using dynamic difficulty adaption, we used the game's inherent avatar structure to select avatars whose strengths align with the user profile. This user profile is based on a physical activity questionnaire and is not absolute but relative to the group of players. The focus group aims to understand the perception of these recommended avatars regarding choice satisfaction, anticipated level of physical activity, and perceived fairness regarding the co-player's avatars.

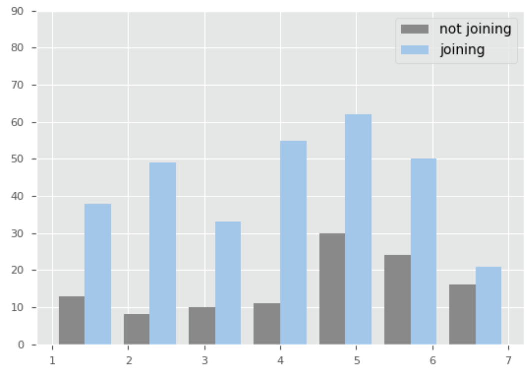
[2]Mario Tennis Aces: https://en.wikipedia.org/wiki/Mario_Tennis_Aces (accessed on 18.05.2023)



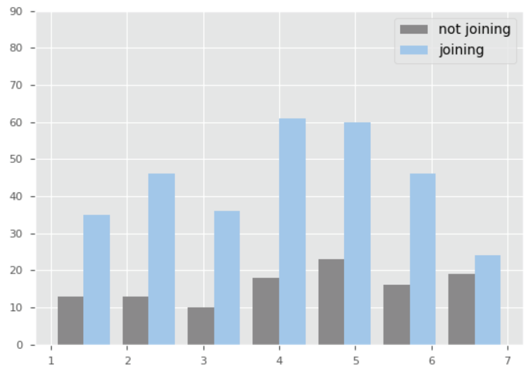
(a) Perceived motivation boost depending on sports



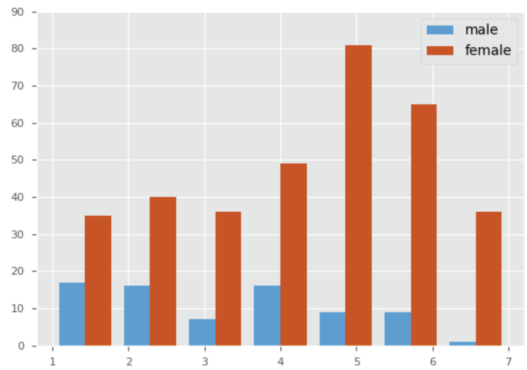
(b) Perceived fairness depending on sports



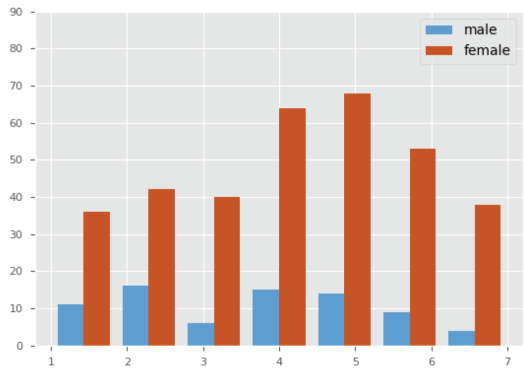
(c) Perceived motivation boost depending on joining



(d) Perceived fairness depending on joining



(e) Perceived motivation boost depending on gender



(f) Perceived fairness depending on gender

Figure 3: These graphs show the differences in perceived motivational boost (left) and fairness (right) of digital adaptivity support in exergames depending on the participants' regular sporting behaviour (top), their willingness to join a group sports activity as the least skilled player (middle) and their gender (bottom).

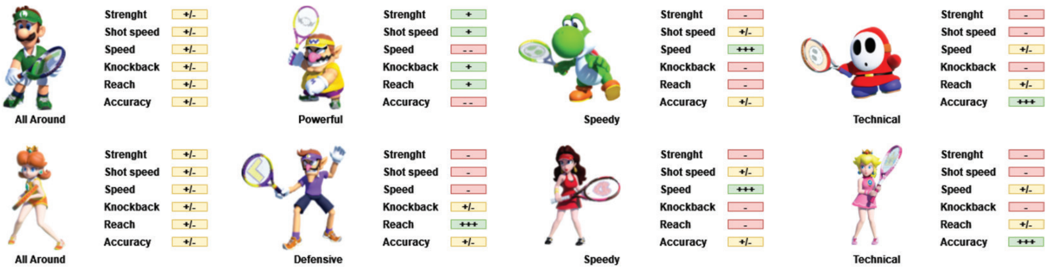


Figure 4: Characters from Mario Tennis Aces that were used in the recommender system with their strengths and weaknesses.

4 FOCUS GROUP ON AVATAR RECOMMENDATIONS

4.1 Study Procedure

After signing informed consent for recording the focus group, the participants were shown trailers of the game and the physical activity option within the game ("Swing Mode"). Then, the participants filled in a questionnaire on their demographics, their physical activity level, and their preferences regarding the eight available avatars. The International Physical Activity Questionnaire (IPAQ) was used to retrieve the participant's physical activity in the previous seven days [6]. Afterwards, a group recommender system was used to link the physical characteristics of the users to the avatar characteristics in the game, recommending the best representative avatars for the group. Once the recommendation was presented to the participants, their first reaction to their recommended avatars was recorded. The participants were further asked how they expected their physical activity and gaming experience to be with this avatar. Then, with the help of a visual overview (Figure 4), the difference in the ability of the various recommended avatars was shown to the participants. This helped them understand what the strengths and weaknesses of each avatar were. Once the background information was revealed, the participants were asked about the perceived social fairness and whether they expected that this recommendation would influence their physical activity during the group gameplay. During the focus group, no active gameplay was involved.

4.2 Avatar Recommender System

We used a well-known commercially available game as the basis for our focus group. The advantage of the game Mario Tennis Aces is that it already provides tennis player avatars with different strength, speed, and technical skills. We used these profiles and the information on the participant's physical activity levels to provide a group recommendation with four avatars. The recommendation was limited to eight different characters in the game, which are shown in Figure 4. Most classes (All-Around, Speedy, Technical) are represented by two avatars in the selected subset, one male and one female. The exception are the Powerful and the Defensive avatar, which are both male and only represented once in this study. This selection was based on a diverse set of strength profiles, while sticking to well-known and recognizable characters. The character of Mario was excluded for having a strong potential positive bias.

The recommender system iteratively tries to match the physical activity characteristics of each participant to the characteristics of the eight avatars. Random draws are used to balance the chances of each participant being matched first. If no solution is found, the matching criteria are relaxed, and the process is reiterated. If, during the iterations, multiple matches are available for the same participant, the ones from earlier/stricter iterations are preferred. The matching of characteristics is based on values from the IPAQ [6], namely minutes of vigorous, moderate, walking, and sitting activities in the past seven days. The avatars are matched to the relative position of players in each activity category of the IPAQ. This also means that the same player could have a different avatar when playing with a different group. The representation of avatars in terms of relative IPAQ data is as follows:

- The Powerful avatar is matched with high vigorous activity and low moderate or high sitting activity.
- The Defensive avatar is matched with high walking activity and low vigorous or low moderate or high sitting activity.
- The Speedy avatars are matched with low sitting and/or high moderate activity and low vigorous activity.
- The Technical avatars are matched with high walking activity and low vigorous activity.
- The all-around avatars are used if no other match is possible.

This representation of physical strengths and weaknesses is mostly perceptual since the game mechanics only weakly represent the actual physical exertion of the player. While this algorithm is only a preliminary approximation, it is easily explainable to the focus group participants and triggers clear associations. In table 1, these iterations are demonstrated for the four participants of the focus group. Based on the definition of avatar classes in terms of IPAQ strengths and weaknesses, the recommendation algorithm takes the following steps to determine the group's ideal avatar distribution.

(1) Determine a relative ranking of the group's members for each IPAQ category.
(2) For each avatar class, randomly check for each group member whether they match in relative strengths/ranking in the respective IPAQ categories (first come, first served).
(3) If user does not match any class, assign the All-Around class
(4) Repeat the matching with decreasing strictness (i.e. not just highest/lowest rank) until each player has been assigned a non-All-Around class.
(5) Select the final avatar class based on the earliest possible iteration while avoiding two players receiving the same avatar
(6) If there are two avatars in one class, use the user's preference to choose between them.

Table 1: Participant's minutes in each activity category of the IPAQ questionnaire and their resulting avatar recommendations.

ID	Age	Vigorous	Moderate	Walk	Sitting	1 st Ranking	2nd Ranking	3rd Ranking	Final Avatar
1	23	300	500	495	1200	All-Around	Defensive	Powerful, Defensive, Technical	Waluigi (Defensive)
2	22	480	400	52	1260	All-Around	All-Around	Powerful, Defensive	Luigi (All-Around)
3	23	360	960	1080	1980	All-Around	Speedy	Speedy	Yoshi (Speedy)
4	23	630	1120	200	1380	All-Around	Powerful, Defensive	Powerful, Defensive	Wario (Powerful)

4.3 Participants

The four participants in the focus group were recruited using convenience sampling. Since the game only allows groups of four players to compete in a two-versus-two tennis match, the recommender system and the focus group also target four participants. While the game was not played during the focus group, the general criteria for participation were still incorporated, namely some experience with video games and no impairment to physical activity. In the final set of focus group participants, all four participants were male, and they knew each other before participating. The average reported age was 22.8 (SD=0.4). The physical activity level of the participants in the different categories and the resulting recommendations in consecutive iterations of the algorithm are shown in Table 1.

4.4 Results

The first question after receiving the recommended avatars is aimed at the initial responses of the different participants. One participant was very happy with his recommendation due to the fact that this was coincidentally also his preferred avatar. The other three participants had hoped for a different recommendation and wanted to trade their avatars with other participants in the group. When asked about their expected experience (physical and gameplay) with their recommended avatar, responses were generally positive. Two participants expected no major difference in experience to the avatars they would choose when playing the game, while the other two were either curious or excited to test their avatars. After explaining the avatar recommendation and the strengths and weaknesses of the avatars, the participants were asked for their opinion on the fairness of the recommendation. Two participants expected to have an advantage by using the strengths of their characters and the weaknesses of the opponent's characters. However, all participants expected the game to be fair, as Mario games tend to be well-balanced. Regarding their expected physical activity during gameplay, they all agreed that knowing the reasons for the mapping and their avatar's abilities would make them try to use its strengths.

5 DISCUSSION AND LIMITATIONS

In this section, we discuss (1) the insights we gained on how players adapt in group sports, (2) who would like this adaption to be supported digitally and (3) how users reacted to a group-adapted avatar recommender system that was based on their physical activity levels. Finally, we discuss the limitations of this work.

5.1 Human Adaptivity in Group Sports

It becomes clear from the survey results that participants in group sports activities are generally very willing to adapt their own behaviour to other inferior co-players if needed. This is an important finding towards the acceptance of such adaptations in an exergame by the superior players, for whom the adaptation would pose a disadvantage. However, we did not investigate whether there is a limit to the strength of adaptation they are willing to accept. Further, despite the high willingness to adapt, fewer people would expect others to adapt to them, leading to less motivation for joining sports activities with superior co-players, especially for non-sporty participants. This dilemma could be avoided by the exergame acting as an impartial balancing party, without the need for asking or expecting others to adapt their behaviour.

5.2 Potential Target Groups

The survey further shows that certain groups of people are more likely to wish for digital support regarding group adaptivity than others. Based on the willingness to join as inferior co-players, especially non-sporty people would be more likely to join the activity if they felt less inferior. Their perception of the support was not significantly more positive than the perception by other participants, but when limiting the group to those players who would actually not join the activities as inferior players, digital support is perceived as significantly more motivational than by others. Finally, it seems that female participants would expect greater benefits and fairness from such a system than male participants. This is in line with previous research on gamification and persuasion, where females tend to prefer collaborative rather than competitive strategies [25].

5.3 Perception of Initial Recommender

The general perception of an initial group recommender system for avatars in Mario Tennis Aces was positive. However, the comments from the group led to interesting considerations for future iterations of such systems. For example, all participants valued that the game was balanced independently of their avatar choices, which highly influenced their perception of fairness. As such, the group-adaptive support would then mostly go towards emphasizing or balancing specific strengths and weaknesses rather than giving overall advantages to inferior players through lowered game difficulty. This would indicate that already small difficulty changes could trigger increased motivation. Further, the proposed avatar was initially not liked by most participants and only accepted and embraced after an explanation of the underlying reasoning. This indicates a need for explainable recommendations, which in turn might lead to privacy considerations concerning the actual physical fitness and health of participants. It further confirms the idea that identification is an important part of motivation through avatar choices. In the future, this could also be done by choosing more realistic personalized avatars or considering additional preferences (e.g. gender, character) in the recommendation.

5.4 Limitations

We must report an obvious limitation regarding the bias in our survey sample towards sporty female participants, while our focus group was conducted with a convenience sample of only male participants. A further limitation is the lack of actual experience of an exergame situation during both the survey and the focus group session, as participants only reflected on the avatars but did not play the game with their respectively assigned avatars. Exploring such a connection to actual mental and physical experiences will be very valuable in future studies.

6 CONCLUSION AND FUTURE WORK

In this paper, we present two studies on the use of group adaptivity in exergames to foster more physical activity. A survey among 421 participants revealed that people are willing to adapt their own behaviour in a group sports situation but do not expect others to do so. They are further less likely to join a group sports activity if they feel less skilled than the other participants. Digital support for group adaptivity in exergames received mixed reactions. However, the group that would not join the activity if less skilled consider such support more likely to foster their participation compared to people that would join the activity anyways. Finally, female participants perceive digital support for group adaptivity as fairer and more likely to foster their participation than male participants. In the second study, a focus group was presented with avatars for an exergame, Mario Tennis Aces, that were recommended based on their physical activity patterns in the past seven days. The focus group displayed a generally positive attitude toward the avatars, the expected experience, and perceived fairness once the process of matching was revealed. Overall, these results show promise for the future implementation and use of digital support for group adaptivity in exergames, especially for people who do not regularly practice sports, who do not want to join games with superior co-players, or female players. In the future, we plan to conduct studies in a similar setup as the focus group while additionally measuring heart rate and satisfaction during actual gameplay and comparing self-selected avatars to recommended avatars. If individual sessions show a positive impact of the adaptivity on physical activity during a gaming session, a future follow-up could be a real-life study of group sports or exergame activities with an experience sampling methodology [21] for a more longitudinal perspective.

REFERENCES

[1] Alexandro Andrade, Clara Knierim Correia, and Danilo Reis Coimbra. 2019. The psychological effects of exergames for children and adolescents with obesity: a systematic review and meta-analysis. *Cyberpsychology, Behavior, and Social Networking* 22, 11 (2019), 724–735.

[2] Christiane Attig and Thomas Franke. 2020. Abandonment of personal quantification: A review and empirical study investigating reasons for wearable activity tracking attrition. *Computers in Human Behavior* 102 (2020), 223–237.

[3] Franziska Baum, Elisabeth Vogel, Christiane Attig, Martina Kanning, Tiare Feuchtner, Hanna Hauptmann, Ulf-Dietrich Reips, Thomas Franke, and Laura M König. 2022. The what and the why of using digital interventions for promoting physical activity and healthy eating: A quantitative online survey. https://osf.io/f7nrx?view_only=b280c9857d644e7cad46957d6bbf2968

[4] Aslihan Tece Bayrak, Rahul Kumar, Jak Tan, DeVon AhMu, Jordan Hohepa, Lindsay A Shaw, Christof Lutteroth, and Burkhard C Wünsche. 2017. Balancing different fitness levels in competitive exergames based on heart rate and performance. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction*. 210–217.

[5] J Cohen. 1992. A power primer Psychological Bulletin 112: 155–159. *IE and Dyadic Adjustment* 1 (1992).

[6] IPAQ Research Committee et al. 2005. International physical activity questionnaire: Short last 7 days self-administered format. *http://www.ipaq.ki.se* (2005).

[7] Amy V Creaser, Stacy A Clemes, Silvia Costa, Jennifer Hall, Nicola D Ridgers, Sally E Barber, and Daniel D Bingham. 2021. The acceptability, feasibility, and effectiveness of wearable activity trackers for increasing physical activity in children and adolescents: a systematic review. *International journal of environmental research and public health* 18, 12 (2021), 6211.

[8] Samuel T Forlenza, Norbert L Kerr, Brandon C Irwin, and Deborah L Feltz. 2012. Is my exercise partner similar enough? Partner characteristics as a moderator of the Köhler effect in exergames. *GAMES FOR HEALTH: Research, Development, and Clinical Applications* 1, 6 (2012), 436–441.

[9] Shaun Gallagher. 1986. Body image and body schema: A conceptual clarification. *The Journal of mind and behavior* (1986), 541–554.

[10] Frederick X Gibbons and Bram P Buunk. 1999. Individual differences in social comparison: development of a scale of social comparison orientation. *Journal of personality and social psychology* 76, 1 (1999), 129.

[11] Appnori Inc. 2018. Squash Kings VR. Retrieved May 19, 2023 from https://store.steampowered.com/app/787460/Squash_Kings_VR/

[12] Konstantina Kilteni, Raphaela Groten, and Mel Slater. 2012. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments* 21, 4 (2012), 373–387.

[13] Laura M König, Gudrun Sproesser, Harald T Schupp, and Britta Renner. 2018. Describing the process of adopting nutrition and fitness apps: behavior stage model approach. *JMIR mHealth and uHealth* 6, 3 (2018), e8261.

[14] For Fun Labs. 2016. Eleven VR. Retrieved May 19, 2023 from https://elevenvr.com/

[15] Judith Masthoff. 2015. Group recommender systems: aggregation, satisfaction and group attributes. *recommender systems handbook* (2015), 743–776.

[16] Judith Masthoff and Albert Gatt. 2006. In pursuit of satisfaction and the prevention of embarrassment: affective state in group recommender systems. *User Modeling and User-Adapted Interaction* 16 (2006), 281–319.

[17] Amir Matallaoui, Jonna Koivisto, Juho Hamari, and Ruediger Zarnekow. 2017. How effective is "exergamification"? A systematic review on the effectiveness of gamification features in exergames. (2017).

[18] Ivelina V Piryankova, Hong Yu Wong, Sally A Linkenauger, Catherine Stinson, Matthew R Longo, Heinrich H Bülthoff, and Betty J Mohler. 2014. Owning an overweight or underweight body: distinguishing the physical, experienced and virtual body. *PloS one* 9, 8 (2014), e103428.

[19] Beatrice Rammstedt and Oliver P John. 2007. Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of research in Personality* 41, 1 (2007), 203–212.

[20] Simone M Schneider and Jürgen Schupp. 2014. Individual differences in social comparison and its consequences for life satisfaction: introducing a short scale of the Iowa–Netherlands Comparison Orientation Measure. *Social Indicators Research* 115 (2014), 767–789.

[21] Yury Shevchenko, Tim Kuhlmann, and Ulf-Dietrich Reips. 2021. Samply: A user-friendly smartphone app and web-based means of scheduling and sending mobile notifications for experience-sampling research. *Behavior Research Methods* 53, 4 (2021), 1710–1730.

[22] Erik Sikström, Amalia De Götzen, and Stefania Serafin. 2015. Self-characterstics and sound in immersive virtual reality—Estimating avatar weight from footstep sounds. In *2015 IEEE Virtual Reality (VR)*. IEEE, 283–284.

[23] Ricardo Borges Viana, Vinnycius Nunes de Oliveira, Scott J Dankel, Jeremy P Loenneke, Takashi Abe, Wellington Fernando da Silva, Naiane Silva Morais, Rodrigo Luiz Vancini, Marília Santos Andrade, and Claudio Andre Barbosa de Lira. 2021. The effects of exergames on muscle strength: A systematic review and meta-analysis. *Scandinavian Journal of Medicine & Science in Sports* 31, 8 (2021), 1592–1611.

[24] Karoline Villinger, Deborah R Wahl, Heiner Boeing, Harald T Schupp, and Britta Renner. 2019. The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: A systematic review and meta-analysis. *Obesity reviews* 20, 10 (2019), 1465–1484.

[25] Mark Van Vugt, David de Cremer, and Dirk P Janssen. 2007. Gender differences in cooperation and competition: The male-warrior hypothesis. *Psychological science* 18, 1 (2007), 19–23.

[26] Wenge Xu, Hai-Ning Liang, Zeying Zhang, and Nilufar Baghaei. 2020. Studying the effect of display type and viewing perspective on user experience in virtual reality exergames. *Games for health journal* 9, 6 (2020), 405–414.

[27] Soojeong Yoo, Christopher Ackad, Tristan Heywood, and Judy Kay. 2017. Evaluating the actual and perceived exertion provided by virtual reality games. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 3050–3057.