Adaptive Difficulty in Exergames for Parkinson's disease Patients

Jan Smeddinck*
TZI - Digital Media Research Group
University of Bremen

Sandra Siegel**
TZI - Digital Media Research Group
University of Bremen

Marc Herrlich***
TZI - Digital Media Research Group
University of Bremen

ABSTRACT

Parkinson's disease (PD) patients can benefit from regular physical exercises which may ease their symptoms and can slow down the progression of the disease. Motion-based video games can provide motivation to carry out the often repetitive exercises, as long as they establish a suitable balance between the level of difficulty and each player's skills. We present an adaptive game system concept, which is based on separate difficulty parameters for speed, accuracy and range of motion. We then describe the heuristic performance-evaluation and adjustment mechanisms in a prototypical implementation which was applied in a case study with three PD patients over a period of three weeks. Results indicate that the system facilitated a challenging yet suitable game experience and a detailed analysis of the results informed a number of follow-up research questions for future research.

Keywords: Older adults, serious games, motion-based games, dynamic difficulty adjustments, adaptive systems, personalization.

Index Terms: K.4.2 [Computers and Society]: Social Issues – *Assistive technologies for people with disabilities*; K.8.0 [Personal Computing]: General – *Games*.

1 INTRODUCTION

Parkinson's disease (PD) is a neurological disease with a variety of motor and non-motor symptoms. Regular physical exercise plays an important role in slowing down progression of the disease. Such exercise is usually instructed in physical therapy sessions, but patients are also encouraged to continue exercising at home [1]. Recently, video games have been identified as a possible means of motivating patients to perform the often repetitive exercises [2]. However, a range of challenges remain on the path to creating games that can truly be used as kinesiatric digital games (KDGs) by players from strongly heterogeneous target groups such as PD patients. A major challenge that is frequently mentioned in related literature is the need to provide adjustable levels of difficulty, depending on age, state of disease and the intensity and types of symptoms [2]–[6]. Keeping the level of difficulty in balance with the players' capabilities and limitations is not a one-off task, since human factors vary both in the long-term and the short-term. This results in a need for automatic adaptive difficulty systems that do not rely on frequent manual interventions to perform adaptations. This study presents an analysis of the problem space of adaptive difficulty for KDGs. An adaptive game system concept with a focus on generally applicable game difficulty parameters and a prototypical implementation that were devised from this analysis are described in detail. Finally, we present a user study in which the system was observed in action with three subjects over a period of five play sessions that were scattered over three weeks. We took a formative triangulation approach based on observations, questionnaires, interviews, the analysis of game logs and a post-hoc analysis by therapists. The study aimed at exploring the suitability of the system to adapt difficulty settings to be appropriate in terms of the resulting physical exercises and to generate a good game experience for each individual subject. Finally, a number of informed research questions for future research in the area of adaptive difficulty for KDGs are devised based on observations from the case study.

*smeddinck@tzi.de, **siegel@tzi.de, ***mh@tzi.de

2 KINESIATRIC DIGITAL GAMES FOR PHYSICAL THERAPY AND PARKINSON'S DISEASE

In general terms, a lack of dopamine in PD patients causes motor disorders that affect gait, the ability to perform everyday activities and communication skills. The most common symptoms are tremor at rest, rigidity and *bradykinesia*, the latter being the most characteristic and disabling. It refers to a reduced range of motion (*hypokinesia*) and the (sudden) inability to initiate and execute movements (*akinesia*). PD also affects fine motor skills and other functions, such as swallowing, speech and facial expressions. Besides medication, physical therapy is recommended at all stages in order to slow down the progression of the disease [1]. Full-body motion-based games have the potential to slow down cognitive decline [7] and to support physical therapy [4], [6]. While mass market games for full-body tracking systems, such as the Kinect[1] are currently very popular, the motor skills they require are often beyond the skills of heterogeneous target groups that could benefit from games for kinesiatric purposes [2], [6], [8]. Accordingly, a growing number of research projects are currently exploring the design and potential application scenarios of KDGs that match the specific requirements of a range of heterogeneous target groups [5], [6], [8], [9]. So far, the target group of Parkinson's disease patients has not received much attention [2].

2.1 Adaptive Difficulty

Ideally, a player of KDGs would find the games enjoyable and the game movements pleasurable, while not noticing the repetitive nature of the exercises or feeling physically overstrained. Video games can put players into a state of flow, where they are lost in the experience of the moment that can arguably facilitate a positive game experience as mentioned above. According to Csikszentmihalyi [10], activities with clear goals that are challenging, require skills, provide immediate feedback and can be completed may enable flow. These prerequisites clearly demand a balance of the level of challenge that a game presents and the players' skills. Notably, the key enabling factor is not the relation of objective game difficulty to objective player skills, but the players' perception of their own competence [10]. As Sinclair

[1] http://kinectforwindows.org/

et al. point out in their model of “dual-flow” for motion-based video games [11], the balance of physical challenges and fitness needs to be considered in addition to the balance related to cognitive and fine-motor skills.

Adequate difficulty settings are usually the result of extensive play testing [11] and in most games, personalization options concerning the level of difficulty are very limited. They can, for example, be adjusted by the means of a simple combined multi-tier parameter that represents “easy”, “medium” and “hard” difficulty, or by controlling the amount of available game resources. More recently a number of games implement *Dynamic Difficulty Adjustments* (DDA), where individual players’ abilities are detected and game settings are automatically adapted [12]. The system we propose herein follows the line of heuristics-based DDA. However, it provides a level of adaptability, or means for therapists to intervene, which has not been a common focus of explorations of DDA.

We argue that, in order to be useful, safe and fun, KDGs require not only mechanisms that facilitate fine-grained manual control of the level of difficulty (making them *adaptable*) [6], [13]. They also require mechanisms that automate the on-going adjustments of the parameters that define game difficulty, based on the performance and game experience of the players (making them *adaptive*). This requirement was confirmed by therapists in a semi-structured interview during a pre-study. Both interviewed therapists said that fine grained manual control would be desirable. However, they also pointed out that frequent manual adjustments for every individual patient would require valuable time they would rather spend with physical therapy sessions.

Existing research projects regarding KDGs have isolated three commonly applicable generic difficulty parameters: required *accuracy* [6], range of motion (*amplitude*) and *speed* [3], [6], [13]. While some projects have experimented with DDA for speed and accuracy settings, the amplitude settings are usually adjusted by the means of a (repetitive) calibration procedure [3], [6], [13].

3 Adaptive Game System Concept

Since adaptive KDGs are to be operated by therapists (and potentially also by the players themselves), the detailed technical game difficulty parameters that are implemented in the game mechanics need to be translated into a comprehensive set of generic difficulty settings. Related work indicates that games for PD patients should focus on continuous large and fluent movements [14] and found that the players often do not fully exploit their potential amplitude if they are not specifically encouraged to do so [2]. Thus, one important generic setting is concerned with the amplitude that the game will demand its players to perform. Control over this can either be achieved via the placement of items or hit-areas, or by scaling the input modalities. Often, the latter method is preferable since being able to reach items that are placed far-out is an important motivational factor [6]. In addition to the amplitude setting, settings to control the game speed (e.g. via allowed reaction times) and the required accuracy (e.g. via the size of hit-boxes or via the tolerance of gesture detection) are necessary to adapt games for PD patients, because they relate to typical symptoms of PD (e.g. *bradykinesia*). Further parameters that frequently play a role in adapting KDGs, such as endurance, cognitive complexity and cognitive resilience are omitted in this description, since they do not play an important role for the specific game employed in this study.

3.1 Performance-based Adaptivity

Adaptive games require two basic components: a *performance-evaluation* and an *adjustment mechanism* [15]. Their concrete implementation may range from simple heuristics to AI-based approaches for complex games [12]. In the case of KDGs for PD patients the games that resulted from an iterative user-centered design process [2] are often rather well-defined casual games. In such cases, heuristics, while needing to be tailored specifically for each game, can offer more control over both mechanisms. In order to facilitate adjustments for each individual player, sets of player profiles, their current settings, and the concerned games need to be stored. In KDGs for heterogeneous audiences, settings may also need to be differentiated depending on different body parts, e.g. to adjust to typical asymmetries [1] in PD symptoms.

Important design decisions when creating adaptive games are to determine the point in time and the frequency of adjustments to the settings. Frequent small-scale settings that are applied during on-going play sessions may more closely approximate optimal settings for an individual player at a given point in time. However, in addition to the risk of overfitting, when adjusting amplitude settings via the scaling of input controls, changes during play pose a serious risk of breaking immersion. Thus, in the system detailed below, adaptations are only applied between game rounds. To facilitate warm-up (an important element of physical therapy or exercise sessions), the current difficulty setting for each player is reduced by a small amount after considerable breaks.

When players first begin playing a game, performance data to drive the adjustment mechanism for the settings is not yet available. Standard medical assessment scales have the potential to inform well matching initial settings for a player [16]. In practical use-cases, however, such data is not always available and the procedures are often time-intensive and involve private medical data. Thus, in a more generally applicable approach, it is possible to rely on rather low default settings (that render the game easily playable for a large majority of players from the target group), a calibration process, where necessary, and on aiming for a swift initial adaptivity that nears the projected optimal settings in initially large steps.

3.2 Calibration

In the study at hand, amplitude plays a central role and an initial calibration was required to determine the players’ maximum distance of reach, which is dependent on their agility and physical proportions, in order to establish constraints that help to prevent overstraining or injuries [17]. The process should be as fast and unintrusive as possible. It should also deliver results in the required accuracy and facilitate some manual control for therapists to influence the constraints based on their personal knowledge about the individual players. Games can also implement reminders for therapists to repeat calibration when setting constraints are reached.

4 Prototype Implementation

The game implemented for this study is called *Sterntaler*. Players use their hands to collect stars that appear sequentially, following a predefined set of paths that reflect physical therapy movements. When the stars on a path are collected in the order of appearance, a shooting star appears that can be collected for bonus points. The star paths appear mirrored for each hand and are selected randomly. After each round, players are shown a score screen that also contains a motivating message which is always positive but reflects the players’ performance (e.g. “Great! You strongly improved.” or “Super! You’ve done it.”)[2].

[2] All quotations relating to the study presented in this paper are translated from German.

The game is implemented in C#, based on the Microsoft XNA framework, using .NET and WPF. Player tracking is handled with a Kinect device using the official SDK. Player profiles and game logs are saved in a MongoDB data store. The Kinect device delivers a skeleton abstraction of the players' body posture in absolute metric space. As a preprocessing step, this skeleton is translated so that the upper back takes a well-defined central position on screen.

Figure 1: *Sterntaler* screen. Coins drop when stars are fetched.

4.1 Dynamic Difficulty Parameter

The parameters for speed and accuracy are normalized to a range from 0.0 to 1.0 with mapped boundary values resulting from play testing. The amplitude parameter may exceed 1.0, since players may vary greatly in body size. All parameters default to 0.3.

The *speed* parameter defines the time that is available to collect stars on the paths before they disappear. Each star path has a phase in which stars appear that equals the minimal completion time, an optional break, and a phase in which stars disappear. All three durations add up to the maximal completion time. The exact functions, which were determined based on play testing, assure that, with an increasing setting for the speed parameter, the maximal completion time is reduced faster than the minimal completion time. For high speed settings, additional star paths appear to assure that a game round lasts approx. 90 seconds.

Play testing revealed that the *accuracy* parameter does not influence the level of difficulty for *Sterntaler* as strongly as adjustments of the required speed or amplitude. Thus, the parameter, which represents the hit-areas of hands and stars, is simply decreased linearly up to 30% less than the base value as the parameter setting increases.

The *amplitude* parameter is derived from the desired distance of the hands above the shoulder that is needed to reach the topmost star (the setting implies an approximately equidistant amplitude in other directions). It is defined as:

$$amplitudeDifficulty = 2 * (distance_{aboveShoulder} - 0.15) \quad (1)$$

The subtraction of 0.15 induces a minimal amplitude of 15 cm. The setting is then converted into a scale factor for the input data based on pixel distance.

4.2 Performance Metrics

For more straight forward internal and conceptual handling, the performance of players of *Sterntaler* is measured with three heuristics that produce normalized outputs ranging from 0.0 to 1.0.

$$collectingPerformance = \frac{stars_{collected}}{stars_{total}} \quad (2)$$

Low scores on the *collecting performance* indicate that the game was too fast, or that the stars were not reachable.

A low score on the *time performance* indicates that the stars were collected late or not at all. Time performance is defined as one minus the ratio of the time that a player needed to collect all paths to the maximum collection time of all paths.

$$timePerformance = 1 - \frac{\sum_{k=1}^{paths_{total}} \left(\frac{usedTime_k - appearanceTime_k}{totalTime_k - appearanceTime_k}\right)}{paths_{total}} \quad (3)$$

The *amplitude performance* is calculated as the ratio of well reachable stars to overall stars in order to avoid counting stars that were not collected, but within reach.

$$amplitudePerformance = \frac{\sum_{k=1}^{stars_{collected}}(distance_k) + \sum_{k=1}^{stars_{reachable}}(distance_k)}{\sum_{k=1}^{stars_{all}}(distance_k)} \quad (4)$$

Stars are weighted by their distance to the center position that serves as a translation target for the skeleton. A star that was not collected is only counted as out of reach, if it was located in an area that the players' hands had not moved over during the session (see fig. 2). A high score in amplitude performance indicates that the player was able to reach most stars.

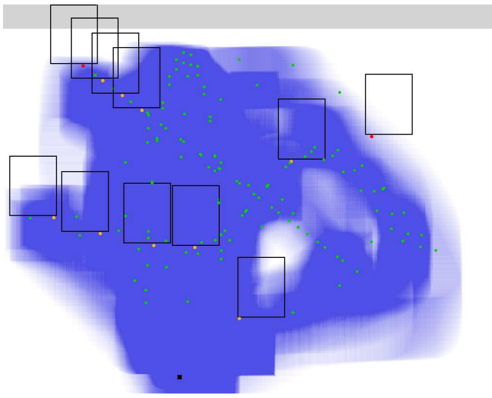

Figure 2: Detecting not-collected and not-reachable stars.

4.3 Adaptation Mechanism

The following adjustment mechanisms and their parameter values were determined through play testing and were only tested for the *Sterntaler* game. In order to avoid overfitting, game parameters can only be adjusted by +/- 0.1 or +/- 0.2 after each round of play. Since accuracy plays a minor role, it is coupled to the speed setting. Both are increased by 0.1 if the collecting performance exceeds 0.93. They are increased by 0.2 if the time performance exceeds 0.8 at the same time. Speed and accuracy are decreased by 0.1 if the collecting performance drops below 0.8, or by 0.2, if the collecting performance drops below 0.7. The thresholds for amplitude adjustments are at 0.95 for and increment and at 0.9 for a decrement. Additionally, a setting can only increase if no other setting is decreased at the same time. In order to facilitate warm-up, the difficulty for next game round is automatically decreased by 0.1 if the player did not play for 3 hours after the last round. If carried out, the calibration procedure takes approximately one minute and overrides the default amplitude parameter setting with a value that is fixed at 0.4 below the amplitude constraint detected during the calibration, since the game movements differ from the rather controlled movements during calibration.

5 CASE STUDY

Validating approaches to adaptive difficulty requires observation over multiple sessions, if the system is to be tested beyond the initial adaptation. Since classical comparative experimental procedures are typically repetitive and require a large number of subjects to investigate very specific research questions, we chose to evaluate the proposed adaptive game system and the prototypical implementation in a case study. The study aims at providing initial insights into the suitability of the game system to adapt the difficulty settings to be appropriate both in terms of the resulting physical exercises and in generating a good game experience for each individual subject.

5.1 Procedure & Methods

Each subject participated in five 30 minute sessions with 15 to 20 minutes of gameplay. Two sessions per week were scheduled with each participant. They were not offered any compensation in return for their participation. The game was projected on a wall. A Kinect device was used for full-body motion-based input and connected to a laptop computer that ran the game. An additional camera was positioned to record the movements of the participants during play from a frontal point of view.

Following a fixed procedure, each session was opened by greeting the participant and asking whether they had carried out any exercises that day. The participants filled out a self-assessment manikin (SAM) [18] to assess their affective state prior to playing the game. They then completed four rounds of the game with either hand and answered a range of questions as detailed below. Sessions were closed with a discussion of the experiences of the day that left room for open feedback. The first session also included a general introduction to the study and the participants' consent to participating. The participants were informed that the study would be about testing and improving game difficulty adjustments. Subsequently, an introductory interview was conducted as detailed below. The first session continued with the initial calibration and proceeded as described above, with the exception that it only included right handed play in order not to overstrain the participants. The last session also differed from the description of sessions provided above. It featured an interview that aimed at capturing feedback and reflections of the participants and included left-handed play only. The following subsections provide an overview of all research methods that were applied in the case study.

5.1.1 Observation

During the course of all sessions the experiment conductor took notes about interesting incidents and players' comments regarding difficulty, or elements of difficulty. These notes were analyzed in conjunction with the methods mentioned below.

5.1.2 Therapists' Analysis

Based on exemplary video sets that showed the participants of the study while playing at low and high amplitude settings with both hands, two therapists were asked to describe their impressions of the subjects' movements in general and to assess their strenuousness and effectiveness. Some scenes were selected based on an interesting coupling of patient behaviors and in-game events. For all other scenes, the therapists could not see the current game state in order to ensure a focus on the quality of the patients' movements.

5.1.3 Questionnaires

Related work suggests that game play and experiment sessions should not be too long [17]. Since the study described herein is designed around a repetitive procedure, the total number of questions that players were to be asked after rounds and sessions had to be limited. Accordingly, after each round, players were asked only one question (Q1): "How did you perceive the level of difficulty in this round?" with the answer options ranging from 1 ("too easy") to 5 ("too hard"). This question was always presented in-game as not to interrupt the flow from one round to another. After each phase of four rounds of play with one hand, participants were asked to complete a SAM (introduced with limit descriptions after Lang et al. [19]) with a five-point scale that was presented as a printout for pointing. They then answered a reduced and adjusted version of the game experience questionnaire (GEQ) [20], taking a focus on the aspects of competence and challenge. Answers were provided by selecting one answer option from a printed scale that ranged from 1 ("I don't agree") to 5 ("I agree"). The statements that were read out loud by the conductor in order to prevent misunderstandings were:

Q2 Challenge: "I felt challenged."
Q3 Competence: "I felt successful."
Q4 Challenge (amplitude): "I found it hard to reach the stars."
Q5 Challenge (speed): "I felt time pressure."

Q4 was not originally part of the GEQ and was added in order to include an item that is explicitly concerned with the amplitude.

5.1.4 Interviews

The semi-structured interviews in the first and last session included both close-ended and open questions. The introductory interview aimed at getting to know the individual participants, at understanding their personal deficits and background, potential sidedness of the disease, and their experience with digital technology. The final interview that was conducted after the last session aimed at gaining a more in-depth impression of the participants' game experience, the training effects, and their perception of difficulty.

5.1.5 Game Log Analysis

In order to underline the preceding mostly qualitative measures with some more quantitative information, game logs were stored for later analysis. The basic information stored by the adaptive game system includes: player id, date, time, sensor configuration, playing hand, applied difficulty settings, star paths (shape, scale, stars with position, creation and removal time, collected-by-player flag and the order of appearance), a record of skeleton frames and summarized performance data (maximum possible score, achieved score, playing time, etc.).

5.2 Participants

The participants for the case study were recruited from a local PD gymnastics training group and they were well known to the therapists involved in the study.

Participant 1 (P1) was a 76 year old male who was diagnosed with PD in 2006. His primary symptom was a flexed spine (*camptocormia*). He was right-handed and his PD symptoms were symmetric. He stated that he engaged in roughly 45 – 90 minutes of physical exercise per week and that he used the computer regularly and liked board games.

Participant 2 (P2) was a 76 year old male who was diagnosed with PD in 2009. His primary symptoms were slowness of walking, writing, and vocal communication. He was naturally left-handed, yet was trained to be right-handed during his childhood. After the onset of PD symptoms his right side was more heavily affected, so he returned to left-hand dominance. He stated that he engages in approx. 45 minutes of physical activities per week, uses the computer for communication and likes board games.

Participant 3 (P3) was a 62 year old male who was diagnosed with PD in 1997. He mentioned the impact of the disease on fine motor skills as his most notable symptom. He self-reportedly participates in about 45 minutes of regular therapy and a large number of different types of sports. He rarely uses computers and was skeptical about video games until he first used KDGs.

5.3 Results

5.3.1 Acceptance of the Game System

The adaptive difficulty mechanism frequently adjusted the difficulty parameters (cf. fig. 3). Yet, the game was understandable and well received. The participants gave the game good fun ratings of 4 to 5 ("Playing the game was fun." answer options 1 - "I don't agree" to 5 – "I agree") and indicated that they would like to continue using the game at least once or twice per week. P3 expressed most enthusiasm and pointed out that he appreciated the perspective of being able to play at home independently of weather conditions. The calibration procedure was completed as planned and patients indicated that they did not feel uncomfortable performing the necessary postures.

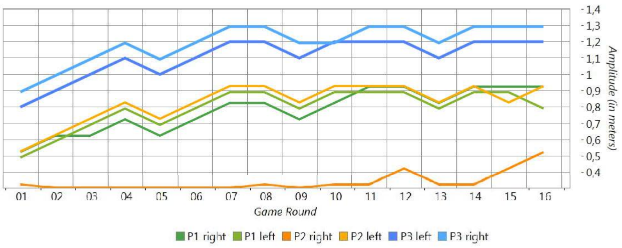

Figure 3: Progression of the amplitude parameter for P1 – P3.

5.3.2 Difficulty Settings and Performance

The initial calibration determined an equal reach for both hands for P1 and P3. P2 had limited amplitude in the right arm and the calibration accordingly determined a difference of 10 cm in vertical stretch (cf. table 1). During the play sessions, speed, accuracy and amplitude settings showed an overall positive trend for all participants (cf. figs. 4, 6, 7, 8).

	Body Height	Right	Left	Both
P1	169cm	46cm/0.62	44cm/0.59	46cm
P2	171cm	36cm/0.42	46cm/0.62	36cm
P3	198cm	64cm/0.90	60cm/0.99	63cm

Table 1. Calibration results: vertical offset shoulder to hands together with the resulting amplitude constraint.

P1 achieved an average collecting performance of 0.89 (SD .05) and an average amplitude performance of 0.97 (SD .02). In round 2-3r (*session 2-round 3, right hand*) when the amplitude setting increased to values close to the calibrated constraint, some difficulties for the player became observable which were due to his flexed spine. In their analysis the therapists validated that, while being challenging, the vertical stretching movements were appropriate considering his symptoms. Due to the time required when getting into an upright position, the speed setting stabilized at a lower level (see fig. 4). In his left-handed play, P1 achieved slightly higher levels with an average collecting performance of 0.9 (SD .09) and an average amplitude performance of 0.97 (SD .04). The log analysis showed that P1 struggled with horizontal movements, since his collection rates for shapes that involved strong horizontal movements was between 83% and 87% compared to more than 90% on average. Therapists criticized that horizontal movements with large amplitude might overstrain this patient and should be limited.

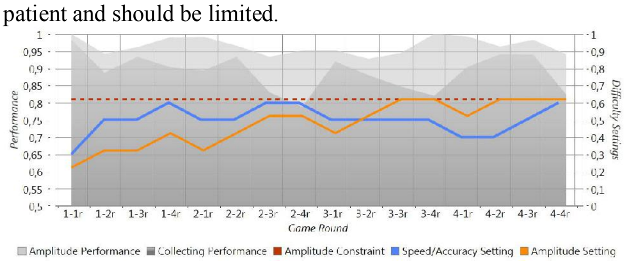

Figure 4: Progression of difficulty and performance for P1 (right).

The performance metrics of P2 showed strong differences between both hands. Based on the calibration procedure, he started with comparatively low amplitude settings. Still, game logs, observation and therapists' analysis revealed that he struggled with collecting stars placed above his head (see fig. 5).

The participant himself agreed that he found it hard to collect these stars when asked in the final interview. Apart from upper stars he had no problem with following the shapes. Accordingly, his speed setting steadily increased. With the right hand, he achieved an average collecting performance of 0.94 (SD .03) and an amplitude performance of 0.95 (SD .04).

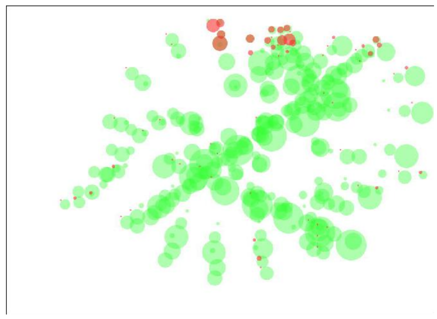

Figure 5: Collected and not-collected stars for P2 (right hand).

His left-handed play did not show strongly notable phenomena and he reached an average collecting performance of 0.9 (SD .05) and an average amplitude performance of 0.95 (SD .04).

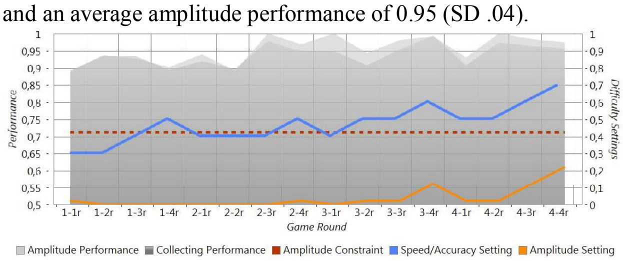

Figure 6: Progression of difficulty and performance for P2 (right).

P3 achieved a high level of collection and amplitude performance throughout all sessions. His average collecting performance was 0.91 for both hands (SD_{right} .08, SD_{left} .05) and his average amplitude performance was 0.97 (SD .04) for the right hand and 0.98 (SD .02) for the left hand.

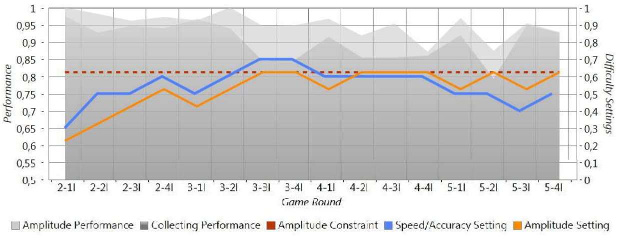

Figure 7: Progression of difficulty and performance for P2 (left).

His performance measures show sharp temporary drops in sessions 3 and 4 (see fig. 8). By observation it became clear that these drops occurred due to tracking errors that were caused by his tall body size of 1.98m. He exceeded the sensor range when taking too many steps forward during play. Furthermore, in the first game rounds of session two and three his performance did not pick up as fast as in the other sessions. These sessions took place early in the morning, when the participant had just taken his medicine. In a post-session clarification he indicated that he did not feel agile until his medicine started to work.

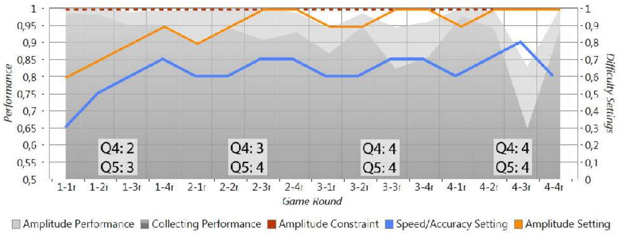

Figure 8: Progression of difficulty and performance for P3 (right).

A Pearson's correlation with a two-tailed significance test for individual cases revealed mostly modest to strong negative correlations between the two difficulty settings and collecting performance, with the exception of the right-handed game rounds of P2 (see table 2).

		Collecting Performance	(right/left)
Speed/ Accuracy	**P1**	r=-0.70 (p<0.01)**	r=-0.60 (p<0.05)*
	P2	r=-0.44 (p=0.9)	r=-0.58 (p<0.05)*
	P3	r=-0.62 (p<0.05)*	r=-0.56 (p<0.05)*
Amplitude	**P1**	r=-0.50 (p<0.05)*	r=--0.4 (p=0.13)
	P2	r=-0.44 (p=0.9)	r=-0.73 (p<0.01)**
	P3	r=-0.51 (p<0.05)*	r=-0.36 (p=0.17)

Table 2. Correlation between difficulty and performance. Significant results indicated by * for p<.05, **for p<.01.

5.3.3 Characteristics & Suitability of Movements

The movements of P1 were characterized by his flexed body posture. For him, standing upright is possible but requires conscious effort. As indicated above, the therapists deemed vertical stretching movements for this patient as demanding but very important. They also pointed out that horizontal movements should be limited for this patient. The frequency and amplitude were otherwise judged as suitable.

For most of the time during play, P2 remained in a static position and reached for stars with his arms and by hip rotation. Vertical arm movements were hindered by problems with his shoulder and his right side was more affected by PD. Notably, the therapists would not have expected the difference in performance to be as strong as it was recorded in the game logs. They assessed the movements to be suitable for this patient and the participant himself found the movements to be pleasant and varied.

Figure 9: Participant 1 while playing.

In one session P2 asked of his own account whether there was a difference between left and right since he perceived "the image to be somehow stretched". Being asked whether this felt unpleasant, he disagreed.

Figure 10:Participant 2 while playing.

P3 performed very dynamic play movements, which involved the whole body, leaning out, and tip-toeing. He self-reportedly appreciated the game movements and even asked for more dynamic game movements, such as jumping. His motion profile involved strong trunk movements with a rather static arm position in low difficulty settings which the therapists assessed as not very appropriate. However, as the adaptive difficulty system quickly adjusted to higher settings, his movements were deemed suitable by therapists. His disease is somewhat asymmetric and he reported that his left arm felt clumsier. Notably, he sometimes supported his weak arm with the other arm. The physical therapists explained that this could be a compensatory strategy to enhance movement control.

Figure 11:Participant 3 while playing.

5.3.4 Perceived Difficulty

The evaluation of perceived difficulty is essential, since a suitable and challenging game experience is not determined by objective difficulty and objective player skills but depends on how players perceive difficulty in relation to their own skills [1].

Question Q1, which was asked after every game round aimed at accumulating the general perceived difficulty. Since participants had problems fixing their rating to a concrete value, half-steps were allowed. The ratings range between 2 and 5. P1 rated the difficulty 3.77 (SD .66) on average, while P2 rated it 3.44 (SD .47) on average and P3 ranked it at 3.45 (SD .66). Thus, general difficulty was perceived as medium to difficult. There was no consistent correlation between collecting performance and perceived difficulty.

Some notable observations were that P1 ranked the game as 5 ("too hard") in rounds 3-3r and 3-4r coinciding with him commenting on the fact that he only managed to collect few shooting stars in these sessions. In accordance with his strong overall performance, P3 was the only participant who provided perceived difficulty rankings below 3. The game rounds he rated with a difficulty of 4 to 5 and 5 were the ones with tracking problems. According to their remarks, the changes in the speed settings were more perceivable to the participants than the changes in amplitude settings.

5.3.5 Challenge & Success

The participants provided generally low assessments of their own success (see table 3). Their comments in the interviews revealed that they related the lack of success to their own skills rather than to the game difficulty, while P3 also attributed his perceived lack of success to his disease. All participants noted that it is important for them to feel successful and they also noted when they performed exceptionally well. For example, the success rankings increased to 4 when P2 achieved his preliminary high-score in session 2, or when P1 managed to collect all shooting stars (commenting: "Yes, that was good!").

	P1	P2	P3
Challenge	4.13 (SD .64)	3.625 (SD .52)	4.125 (SD .35)
Success	2.50 (SD .76)	2.5 (SD .76)	3.37 (SD .74)

Table 3. Average ratings on challenge and success.

However, when questioned in the final interview whether not feeling very successful had a negative effect on their general mood, the participants disagreed. P1 mentioned "It is just a game", P2 responded "No, not at all" and P3 explained "That was just in the game. It does not affect my general mood". This finding corresponds well with the fact that no consistent effects could be detected with the SAM.

The average ratings for challenge by the participants were high, but challenge was conceived as an important element of the game, with P3 noting that the game was "pleasurably difficult", "by no means too hard, it is fun when it's more difficult" and adding that "you only get better when you are challenged". When rating difficulty with 5 ("too hard"), P1 commented: "it should be a challenge". The players were also conscious of the serious background of their gaming activity. Concerning the balance of challenge and success, P3 said: "you only reach a higher level if you get exposed to it" and P2 noted: "it would indeed be nice to be more successful but then it would not serve the purpose".

6 Discussion

As a general trend, the difficulty settings for the three participants stabilized over the later sessions, indicating that the game adapted to the patients capabilities. The adaptive difficulty system was applicable for all three participants and generally determined a challenging but objectively accomplishable level of difficulty. Furthermore, the calibration tool detected appropriate amplitude constraints that were intentionally interpreted rather conservatively. With the exception of P2 in right-handed play, the participants quickly came to play at calibrated constraints or near the constraints. The therapists deemed all movements appropriate with the exception of the extended horizontal movements of P1 and the strong trunk movements of P3 in the first rounds of play at low amplitude settings.

6.1 Follow-up Research Questions

Regarding the small number of participants of this study, generalizing conclusions should be avoided. However, the detailed observations and wide range of methods applied in the case study helped to isolate a number of notable differences between the subjects that clearly highlight the need for even more flexible adaptability and adaptivity. When regarded in detail, these observations can be interpreted as the basis for research questions that should be considered in future work:

The following paragraphs present key observations and central research questions they induce when considered in conjunction with the results of this case study. The questions presented herein aim to be explicit to the use case of KDGs and add to the realm of existing research problems in the larger area of adaptive systems.

With the exception of P2 in right-handed play, all players reached the calibrated amplitude difficulty constraint within four sessions. This finding informs the following research question: *Can a constraint-based adaptive difficulty system estimate more adequate levels of difficulty for players, if the constraints are recalibrated after the difficulty settings have stabilized at near-constraint level?*

While therapists deemed most player movements suitable, they pointed out that P1 should not perform horizontal movements with large amplitude. Future work should thus consider the question: *Can adaptive difficulty systems determine more adequate game settings if more detailed performance measures based on the players' physical capabilities are devised?*

Although the patients mentioned that they wanted to feel successful and achieved high levels of performance, they rated their feeling of success medium to low and their ratings did not correlate to their objective performance. Accordingly, future research should investigate the question: *Which components are essential for designing KDGs that induce a strong feeling of success?*

Therapists indicated that they would appreciate fine-grained control over game settings and they sometimes had specific suggestions concerning which types of movements should be avoided for individual players. The following research questions can be devised: *Which parameters would therapists choose to adapt KDGs and how can games be designed to support these parameters?*

All players in the case study were conscious of the serious aspect of the KDG and underlined that they therefore were ready to accept a challenging game experience. *Does the extrinsic aspect of motivation that represents the serious element of KDGs undermine or extend the effects of intrinsic motivation that is known to be a key factor of flow and positive game experience?*

While playing the game, a participant suspected that some adaptations had been made to his range of motion from one round to the next. *Does being informed about the presence of an adaptive game system notably affect the game performance and / or experience?*

While the questions listed above require further research efforts, the game with adaptive difficulty as presented in this study was perceived as engaging and fun by all participants over the course of five sessions and all participants stated that they would like to play regularly in the future.

7 Conclusion

Parkinson's disease patients can benefit from regular exercising and physical therapy. Kinesiatric digital games have been identified as a potential modality to motivate patients to carry out the often repetitive exercises. However, the target group of PD patients and other target groups of KDGs, such as stroke patients, are very heterogeneous. This makes the adaptability of the games, especially in terms of difficulty beyond the common univariant and discrete levels of difficulty (e.g. "easy", "medium" and "hard") a necessity. Advanced control for therapists and potentially also the players themselves is required. At the same time, the system should be automatically adaptive, since adjustments to the level of difficulty are not a one-off task.

We presented an adaptive game system concept that encompasses parameters for speed, accuracy, and amplitude. Afterwards, a prototypical implementation was described together with the specific heuristics for the dynamic difficulty parameters and the collecting performance, time performance and amplitude performance metrics, as well as the actual difficulty adjustments.

The resulting game was tested in a case study with three participants over the course of five sessions in three weeks. Based on observations, therapists' analysis, questionnaires, interviews and game log analysis we found that the speed and accuracy setting, as well as the amplitude setting, changed frequently and showed an overall trend to increase over time. The amplitude performance neared the calibrated constraint in all two-sided cases but one. On average, the participants perceived the games as challenging and did not feel very successful, which self-reportedly did not have a strong impact on their overall mood. Altogether, the inclusion of detailed separate measures of collecting performance, time performance and range of motion performance and their mapping to the explicit implementation of difficulty parameters for speed, accuracy and amplitude facilitated an adaptive difficulty implementation that produced clearly differentiable and personalized results that were deemed suitable by therapists and appreciated by the patients.

The detailed multi-methodological triangulation performed in this study lead to informed research questions that are specific to KDGs and aim to explore aspects of: constraint-based adaptive systems with recalibration, the impact of the level of detail of performance metrics, creating a feeling of success, the extrinsic motivation due to the serious background of the game, and the impact of being conscious of the adaptive behavior of the system.

Next to these specific aspects, broad and challenging problems remain in devising more generalizable and player-centered adaptive difficulty systems and in evaluating their potential to reliably produce improvements in the skills of the patients, or to noticeably slow down the progression of PD when used over larger spans of time.

8 Acknowledgments

We thank all members of the *Wuppdi!* project, the Bremen regional group of the *Deutsche Parkinson Vereinigung e.V.*, associated physical therapists and the participants for their great support, as well as Maria K. Eckstein for her comments on the draft. This work was partially funded by the *Klaus Tschira Foundation (KTS)* via the grad. school *Advances in Digital Media*.

References

[1] Parkinson's Disease Society, The Professional's Guide to PD. London, 2007.

[2] O. Assad, R. Hermann, D. Lilla, B. Mellies, R. Meyer, L. Shevach, S. Siegel, M. Springer, S. Tiemkeo, J. Voges, and others, "Motion-Based Games for Parkinson's Disease Patients," Entertainment Computing–ICEC 2011, pp. 47–58, 2011.

[3] G. Alankus, A. Lazar, M. May, and C. Kelleher, "Towards customizable games for stroke rehabilitation," in *Proceedings of the 28th international conference on Human factors in computing systems*, pp. 2113–2122, 2010.

[4] J. Burke, M. McNeill, D. Charles, P. Morrow, J. Crosbie, and S. McDonough, "Designing engaging, playable games for rehabilitation," 2010.

[5] K. M. Gerling, I. J. Livingston, L. E. Nacke, and R. L. Mandryk, "Full-Body Motion-Based Game Interaction for Older Adults," in *CHI'12: Proceedings of the 30th international conference on Human factors in computing systems*, 2012.

[6] L. Geurts, V. Vanden Abeele, J. Husson, F. Windey, M. Van Overveldt, J. H. Annema, and S. Desmet, "Digital games for physical therapy: fulfilling the need for calibration and adaptation," in *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, pp. 117–124, 2011.

[7] C. Anderson-Hanley, P. J. Arciero, A. M. Brickman, J. P. Nimon, N. Okuma, S. C. Westen, M. E. Merz, B. D. Pence, J. A. Woods, A. F. Kramer, and E. A. Zimmerman, "Exergaming and Older Adult Cognition," *American Journal of Preventive Medicine*, vol. 42, no. 2, pp. 109–119, Feb. 2012.

[8] S. Flynn, P. Palma, and A. Bender, "Feasibility of using the Sony PlayStation 2 gaming platform for an individual poststroke: a case report," *Journal of neurologic physical therapy*, vol. 31, no. 4, p. 180, 2007.

[9] E. Flores, G. Tobon, E. Cavallaro, F. I. Cavallaro, J. C. Perry, and T. Keller, "Improving patient motivation in game development for motor deficit rehabilitation," *Proceedings of the 2008 International Conference in Advances on Computer Entertainment Technology ACE 08*, vol. 7, no. 8, p. 381, 2008.

[10] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. New York: Harper & Row, 1990.

[11] J. Sinclair, P. Hingston, and M. Masek, "Considerations for the design of exergames," in *Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia*, Perth, Australia, pp. 289–295, 2007.

[12] G. Andrade, G. Ramalho, A. S. Gomes, and V. Corruble, "Dynamic game balancing: An evaluation of user satisfaction," pp. 3–8, 2006.

[13] J. W. Burke, M. D. J. McNeill, D. K. Charles, P. J. Morrow, J. H. Crosbie, and S. M. McDonough, "Optimising engagement for stroke rehabilitation using serious games," *Vis Comput*, vol. 25, no. 12, pp. 1085–1099, Aug. 2009.

[14] B. G. Farley and G. F. Koshland, "Training BIG to move faster: the application of the speed–amplitude relation as a rehabilitation strategy for people with Parkinson's disease," *Exp Brain Res*, vol. 167, no. 3, pp. 462–467, Nov. 2005.

[15] E. Adams, *Fundamentals of Game Design*. New Riders, 2010.

[16] M. Ma, M. McNeill, D. Charles, S. McDonough, J. Crosbie, L. Oliver, and C. McGoldrick, "Adaptive virtual reality games for rehabilitation of motor disorders," in *Proceedings of the 4th international conference on Universal access in human-computer interaction: ambient interaction*, Berlin, Heidelberg, pp. 681–690, 2007.

[17] J. Smeddinck, M. Herrlich, M. Krause, K. Gerling, and R. Malaka, "Did They Really Like the Game? – Challenges in Evaluating Exergames with Older Adults," in *CHI Workshop on Game User Research: Exploring Methodologies*, Austin, Texas, USA, 2012.

[18] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[19] P. Lang, M. Bradley, and B. Cuthbert, *International affective picture system (IAPS): Technical manual and affective ratings*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida, 1999.

[20] K. Gerling, F. Schulte, and M. Masuch, "Designing and Evaluating Digital Games for Frail Elderly Persons," in *International Conference on Advances in Computer Entertainment Technology (ACE 2011)*, Lisbon, Portugal, 2011.