Exergames for Physiotherapy and Rehabilitation: A Medium-term Situated Study of Motivational Aspects and Impact on Functional Reach

Jan David Smeddinck, Marc Herrlich, Rainer Malaka
Research Group Digital Media, University of Bremen, TZI
Bibliothekstr. 1, Bremen, Germany
{smeddinck,mh,malaka}@tzi.de

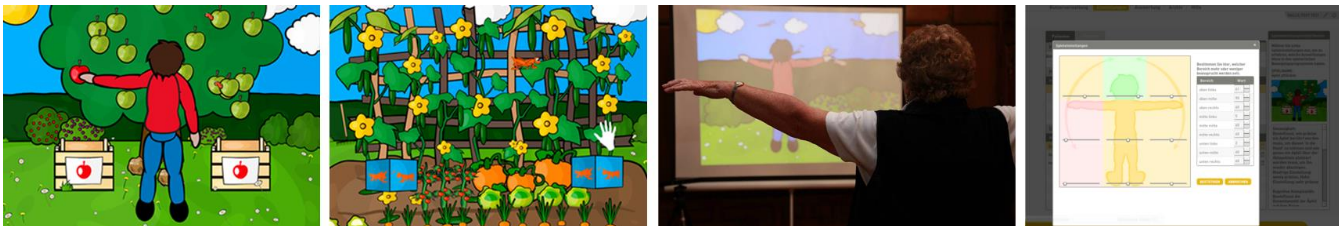

Figure 1: Left to right: screenshots of the games AP, LO, BU (with a player from the target group) and the settings interface for controlling adaptations and adaptivity (settings were defined via individual milestones; the window shows the sub-view for range of motion settings via body zones; all other settings were expressed as percentages and modified via standard WIMP components).

ABSTRACT

Exergames are increasingly considered as an exercise instruction modality in health applications. Studies are typically conducted in non-situated contexts and capture short-term effects. We present first results from a medium-scale study that was conducted over the course of five weeks and integrated into a normal rehabilitation program. The study features three groups, comparing manually adjustable exergames with the identical games in adaptive versions and manual physiotherapy interventions without games. The results indicate that the exergames and traditional therapy are comparable regarding measures of competence and enjoyment, while exergames led to significantly higher scores for autonomy, presence, and in a functional reach test. With traditional therapy, scores for tension-pressure and effort-importance were significantly higher. The initial results of the broader study presented in this paper deliver insights regarding motivational aspects of exergames and traditional therapy and point out which motivational aspects could be strengthened in future implementations.

Author Keywords

Exergames; games for health; player experience; balance.

ACM Classification Keywords

K.8.0 [Personal Computing]: General – Games.

INTRODUCTION

Full-body motion-based games for health (MGH) have emerged as an alternative means for exercise instruction in application scenarios such as physiotherapy (PT) or rehabilitation. Target groups range from the general population, over older adults [2] to specific groups such as people with Parkinson's disease [13] or children with cerebral palsy [5]. All these groups have in common that the heterogeneous abilities and requirements of their members must be met [13]. This requires mechanisms for adaptation that surpass the flexibility of consumer market (e.g. dance/fitness) games. Automatic adaptivity ensues as an additional requirement to lower the burden that frequent manual adjustments place on therapists, doctors or other caregiving staff. Solutions for both challenges, and insights on the general acceptance and effectiveness of MGH, require medium- to long-term studies and evaluations, which are still rare.

We present initial results from a study on the therapy experience, and impact on physical function, comparing a suite of manually adjustable MGH (condition: *normal*) with semi-automatically adaptive versions (*adaptive*) and manual PT interventions without games (*control*), all for upper-body back and neck afflictions. The study was performed in the situated context of a large PT practice with professional therapists (TPs) and real patients. The intervention period lasted for five weeks. The outcomes discussed here are *motivational components*, which were differentiated via measures for the components of self-determination (SDT) [8,10] and *functional therapeutic effects*, which were differentiated via *Timed Up and Go* (TUG) [9] and *Functional Reach* tests (FRT) [7]. We expected the MGH to result in comparable levels of motivation despite less direct personal interaction. The autonomy component of the SDT needs satisfaction model was expected to be increased while others were expected to be decreased with casual game style implementations. Dimensions where the MGH are lacking can hint at future improvements and the dimensions where games are doing well can hint towards design guidelines. We also expected that functional results in the areas targeted by the games can be as good as, or even better than, conventional PT for comparable therapy targets. The initial results of the broader study presented in this paper deliver insights in two areas: How to design for motivation in terms of motivational dimensions and which motivational aspects should be strengthened in future implementations.

STATE OF THE ART

MGH have been used with various target groups in research projects and early clinical applications [1,2,5,13]. The approach of using games for serious purposes in kinesiatric health applications is mostly cited for the potential to provide motivation to execute repetitive exercises [15]. Games (or playful apps) can also offer guidance via immediate custom performance feedback, and objective means of analysis via the collection and processing of performance data [12]. MGH can augment traditional treatments, but they can also allow for improved home/remote exercising. Studies have confirmed that they can be superior to traditional means of exercise instructions, e.g. instruction sheets [14]. While adaptable and adaptive MGH have been designed and evaluated in short-term studies [3,4,13], medium- or long-term situated studies are rare. Yet, these are much needed to confirm the lasting impact and effects. First longer-term studies for MGH indicate positive results (e.g. [5,14]), but real-world situated use is still underreported, esp. in the setting of busy practices, and a detailed motivational dissection has not yet been done. We were motivated to perform such a (rather difficult to organize and control) study since it offers much better comparability with true "in-the-wild" use that cannot be obtained from laboratory studies alone. Therapists and patients have more realistic motives and it is known from related work (and we experienced this in prior studies of our own) that some biases of lab studies are amplified with special target groups such as older adults (e.g. good participant effect). In addition, emotion research with isolated individuals is highly problematic due to the social functions of emotional regulation [6].

SETUP AND PROCEDURE

A suite of three casual MGH was used (Figure 1). Set in a garden scenario, which had been tested to be accepted by a wide range of older adult players, the games were: (a) An apple picking game (AP) which allows players to move at their own speed while picking ripe (red) apples from a tree and placing them in baskets near the ground. (b) A locust catching game (LO) where players were asked to collect locusts from a garden, requiring them to move their hands slowly and in a controlled manner towards the locusts, so as not to scare them away. (c) A butterfly balancing game (BU) that required players to match and hold poses performed by a virtual instructor in order to attract butterflies to sit on the player avatar. The games are part of a larger suite of games developed at the University of Bremen with the support of the "Spiel Dich Fit!" (SDF) project consortium [cf. to ACKNOWLEDGMENTS] over a period of two years. They were created in a player-centered iterative design approach and prior evaluations had shown that they facilitate positive therapy experiences. The selection of games set a focus on upper body flexibility and range of motion relating to the selection criteria for the study into which only patients with recurrent upper-body back and neck afflictions (a common PT target group) were admitted.

The study was performed at a large PT practice in Delmenhorst, Germany led by an experienced and renowned physiotherapist. It included 12 TPs and 29 patients (9 male, 20 female, median age 66 [1st Qu. 59, 3rd Qu. 73]). Patients were split into three groups (*normal* (9), *adaptive* (9) and *control* (11)) by random assignment. A therapist from the same pool was present in each session (S1-5). Patients in all groups were in treatment for chronic spine afflictions and treated following a (pro-active) progressive adaptations therapy approach aiming to reduce pain and to improve flexibility and balance (which continued as usual in the control group and TPs employed the games to continue working towards their goals with the patients in the other groups). Therapists were free to interact with patients as they saw fit in all groups, although they showed less direct interaction with patients in the game groups. They provided frequent verbal feedback and guidance in all sessions. With the games, the patients also received general feedback on their gameplay performance after each round of play but no further feedback on their exercise performance. For the normal group, TPs were regularly reminded that they could use a calibration tool to manipulate personalized settings for *range of motion* (by active game screen zones; cf. Figure 1), required *speed*, *motion accuracy*, *endurance* (level duration) and *complexity* (amount of active objects) for each level. These settings were subject to a pre-study with 10 TPs and were evaluated as being easily understandable, useful for making meaningful adjustments, and allowing for efficient configurations. For the adaptive group, TPs were asked to make initial settings in S1 and to define a target milestone for S5. An automatic rubber-banding adjustment system then aimed at step-wise increasing game settings (in software) while guaranteeing a fair amount of success. In weeks 1 and 3, TPs were reminded that they were always invited to re-adjust the milestones / settings.

Weekly sessions of 20-25 minutes with 10-15 minutes play time per participant (playing each game if in a game group) were held over five consecutive weeks and substituted regular therapy appointments. The games were played in a separate section of a training room at the practice. The measures ranged from affect to exhaustion, self-efficacy, game performance, observational protocols and therapist reports on each patient and have been employed successfully in related work. In this initial report, we focus on two categories of measures collected after S1 and S5 of each participant respectively: (a) *Motivational measures* from instruments rooted in self-determination theory, namely the *competence*, *intuitive control*, *autonomy* and *presence* dimensions of the Player Experience of Needs Satisfaction (PENS) questionnaire [10] and the *tension-pressure*, *interest-enjoyment* and *effort-importance* dimensions of the intrinsic motivation inventory (IMI) [8]. Items of both questionnaires were adjusted to refer to “movement applications” in order to match the settings of all groups. (b) *Functional measures* for evaluating therapeutic effect, namely TUG and FRT. Both categories reflect important aspects of the targeted therapy. SDT motivational measures have been shown to impact exercising adherence [8] and the FRT is a validated indicator for balance and fall prevention [7].

RESULTS

A mixed-model analysis of variances (ANOVA) was conducted to check for statistically significant differences. Independent variables were ‘group’ (between subjects), and ‘session’ (within subjects) as the reported measures were collected after S1 (pre) and S5 (post). Data imputation (max. 1 case per participant on a sub-item in max. 2 participants per measure) was performed using group means. Variances were winsorized [2] (leveling outliers in the top/bottom .1 quantiles to the trim edge values).

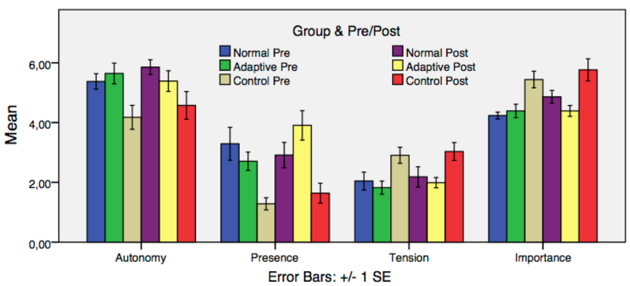

Figure 2: Means and standard errors for autonomy, presence, tension-pressure, and effort-importance collected in the first session (pre) and in the last session after five weeks (post).

ANOVA revealed no statistically significant differences regarding pre and post measures for PENS and IMI. For the independent variable ‘group’ the conducted ANOVA revealed statistically significant differences (Figure 2) for *presence* ($F(2,23)=7.725$; $p=0.003$; partial $\eta^2=0.402$; Levene fail.), *tension* ($F(2,23)=5.584$; $p=0.011$; pt. $\eta^2=0.327$; Lev. pass.), and *effort imp.* ($F(2,23)=7.037$; $p=0.004$; pt. $\eta^2=0.380$; Lev. fail.). The *autonomy* measure very narrowly failed the alpha level for sig. ($p=0.089$) while competence and enjoyment were not significant. Looking at S1 and S5 individually for autonomy revealed statistically sig. diffs. for S1 ($F(2,28)=5.204$; $p=0.013$; Lev. fail.) and a trend ($F(2,25)=3.521$; $p=0.058$; Lev. pass.) for S5. Post-hoc tests confirmed sig. diffs. for *presence* (Dunnett-T3 test used because of failed Lev.; groups *normal* and *control*: $p=0.022$; groups *adaptive* and *control*: $p=0.001$; groups normal and adaptive: n.s.), for *tension* (Bonferroni corrected; groups *normal* and *control*: $p=0.049$; groups *adaptive* and *control*: $p=0.013$; groups *normal* and *adaptive*: n.s.), and for *effort importance* (Dunnett-T3 test used because of failed Lev.; groups *normal* and *control*: n.s.; groups *adaptive* and *control*: $p=0.040$; *normal* and *adaptive*: n.s.).

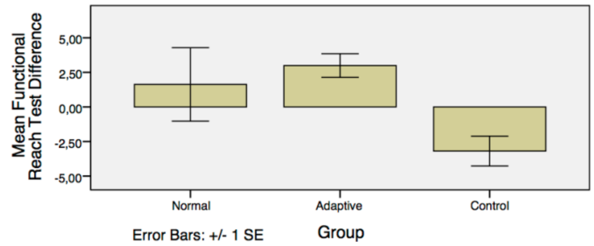

Figure 3: Means and standard errors for the pre-post-differences of the functional reach test for all groups.

An ANOVA for FRT differences (cf. Figure 3) between S1 and S5 revealed statistically sig. diffs. for the indep. var. ‘group’ ($F(2,26)=3.539$; $p=0.045$; pt. $\eta^2=0.228$; Lev. fail.). Post-hoc tests confirmed sig. diffs. for groups *adaptive* and *control* (Dunnett-T3 test used because of failed Lev.; $p=0.001$), all remaining group combinations were not sig.. Table 1 presents the descriptive statistics for measures not shown in Figure 2 or 3 that were comparable across groups.

	Week	*Normal*	*Adaptive*	*Control*
Competence	1	Mean = 5.6 SD = 0.8	Mean = 5.7 SD = 0.8	Mean = 5.6 SD = 0.6
	5	Mean = 5.4 SD = 0.7	Mean = 5.6 SD = 0.8	Mean = 5.7 SD = 0.5
Intuitive control	1	Mean = 6.0 SD = 0.5	Mean = 5.9 SD = 0.7	Mean = 5.2 SD = 0.9
	5	Mean = 6.6 SD = 0.5	Mean = 5.9 SD = 0.6	Mean = 5.4 SD = 1.0
Interest-enjoyment	1	Mean = 6.0 SD = 0.7	Mean = 5.9 SD = 1.0	Mean = 5.6 SD = 0.9
	5	Mean = 6.3 SD = 0.6	Mean = 5.6 SD = 0.9	Mean = 5.3 SD = 1.1
Timed Up and Go Test	W5 - W1	Mean = 0.2 SD = 2.0	Mean =-0.8 SD = 2.1	Mean = 0.2 SD = 3.2

Table 1: Means and std. dev. for competence, intuitive control, interest-enjoyment and the TUG test for weeks 1 and 5.

DISCUSSION

The results indicate that the measures of motivational components are not unanimously in favor of either games or regular PT sessions, pointing to the need to consider motivational aspects of MGH in detail. An explanation for the observed higher mean responses on the tension-pressure dimension in the control condition might be the impact of the strong role of a human instructor. If external pressure induced by the presence of an authority is internalized by the patients (e.g. patients believing that following the lead of their TPs is essential to their own well-being), this does not contradict the increased effort-importance (intrinsic motivation is usually not undermined by internalized extrinsic motivation [11]). Coupling the motivational measures that were significantly in favor of the games conditions could offer an explanation for the observations on these scales. The increased presence in the games conditions is perhaps not surprising and it was also measured with less question items than usual (only two matched the general “movement applications” wording). However, it goes along well with the result of increased autonomy: If a player is getting immersed in a game, it allows them to behave on their own accord. Drawn together, the results suggest that the current generation of games for health (at least the casual style games, which resemble typical applications in this domain) might be a good match for augmenting traditional therapy. Future generations of MGH that are, from the start, designed for patient-therapist and patient-patient interaction through the game might improve results in the other motivational measures and thus facilitate even more captivating exercise instruction. However, partial improvements, or equal performance in MGH already allow for them to be used to augment traditional therapeutic exercise instructions with a new palette of diverse experiences.

Some limitations apply to our study. Our sample shows a gender imbalance. While the split represents the gender distribution in patients of the targeted age group well, further studies are needed to research effects of gender (and age) in more detail. Due to the situated nature of the study and our aim to observe realistic usage, we did not enforce a controlled set of exercises across all groups. The differences we observed can thus only be attributed to the treatment complex of therapy games plus the exercises they implemented as a whole. A more detailed separation requires further (likely lab) studies with a much more controlled approach. Our results also indicate that we cannot clearly separate the *normal* and *adaptive* groups on any measure. While this points to future work, for now the automatic adaptivity does not appear to significantly decrease the PT experience or functional development. At the same time, TPs repeatedly indicated that they appreciated the adaptive version, since it reduces burden on their schedules. The situated context of a busy practice produces a noisy and difficult environment for scientific study. The notable variance in the data we gathered underlines these challenges and the need for improved research methods for situated studies, if more subtle differences are to be detected. Yet, game design for health applications is bound to remain a problem where clear lines are difficult to establish.

CONCLUSION

We presented first results from a medium-term study of full-body motion-based games for physiotherapy and rehabilitation. With a custom-built suite of MGH and a flexible configuration tool, the situated usage in a PT practice was studied over the course of five weeks. Results on the measure of functional reach appear to be increased in the game conditions, hinting at improvements, but further studies are needed in order to generalize this finding. The results of experiential measures rooted in needs-satisfaction models of SDT are mixed and appear to hint at the role of human involvement, indicating directions for future study.

ACKNOWLEDGMENTS

We thank the members of SDF (abraxas Medien Hamm & Wächter GbR and vacances GmbH), Rehamed Delmenhorst (Wolters / Eybe / Schröder-Wolters GbR), Simon Hermsdorf, Sophie Janzen and all participants. The SDF project received funding from the Wirtschaftsförderung Bremen and the EU fund for regional development (EFRE).

REFERENCES

1. Alankus, G., Lazar, A., May, M., and Kelleher, C. Towards customizable games for stroke rehabilitation. *Proc. of the 28th international conference on Human factors in computing systems*, (2010), 2113–2122.
2. Gerling, K.M., Livingston, I.J., Nacke, L.E., and Mandryk, R.L. Full-Body Motion-Based Game Interaction for Older Adults. *CHI'12: Proc. of the 30th int. conf. on Human factors in computing systems*, (2012).
3. Göbel, S., Hardy, S., Wendel, V., Mehm, F., and Steinmetz, R. Serious games for health: personalized exergames. *Proc. of the int. conf. on Multimedia*, ACM (2010), 1663–1666.
4. Hardy, S., Göbel, S., Gutjahr, M., Wiemeyer, J., and Steinmetz, R. Adaptation Model for Indoor Exergames. *Int. Journal of Computer Science in Sport 11*, 1 (2012).
5. Hernandez, H.A., Ye, Z., Graham, T.C., Fehlings, D., and Switzer, L. Designing action-based exergames for children with cerebral palsy. *Proc. of the Conf. on Human Factors in Computing Systems*, (2013), 1261–1270.
6. Kappas, A. Smile when you read this, whether you like it or not: Conceptual challenges to affect detection. *IEEE Trans. on Affective Computing 1*, 1 (2010), 38–41.
7. Katz-Leurer, M., Fisher, I., Neeb, M., Schwartz, I., and Carmeli, E. Reliability and validity of the modified functional reach test at the sub-acute stage post-stroke. *Disability and Rehabilitation 31*, 3 (2009), 243–248.
8. McAuley, E., Duncan, T., and Tammen, V.V. Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: a confirmatory factor analysis. *Research quarterly for exercise and sport 60*, 1 (1989), 48–58.
9. Podsiadlo, D. and Richardson, S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *Journal of the American Geriatrics Society 39*, 2 (1991), 142–148.
10. Rigby, S. and Ryan, R. *The Player Experience of Need Satisfaction(PENS): an applied model and methodology for understanding ke y components of the player experience.* 2007.
11. Ryan, R.M. and Deci, E.L. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology 25*, 1 (2000), 54–67.
12. Smeddinck, J.D., Gerling, K.M., and Malaka, R. Anpassbare Computerspiele für Senioren. *Informatik-Spektrum 37*, 6 (2014), 575–579.
13. Smeddinck, J., Siegel, S., and Herrlich, M. Adaptive Difficulty in Exergames for Parkinson's Disease Patients. *Proc. of the 2013 Graphics Interface Conf.*, Canadian Information Processing Society (2013), 141–148.
14. Uzor, S. and Baillie, L. Investigating the long-term use of exergames in the home with elderly fallers. ACM Press (2014), 2813–2822.
15. Yim, J. and Graham, T.C.N. Using games to increase exercise motivation. *Proceedings of the 2007 conference on Future Play*, ACM (2007), 166–173.