Experimental Gerontology 139 (2020) 111037

Contents lists available at ScienceDirect

Experimental Gerontology

journal homepage: www.elsevier.com/locate/expgero

Balance training monitoring and individual response during unstable vs. stable balance Exergaming in elderly adults: Findings from a randomized controlled trial

Julia Bakker, Lars Donath, Robert Rein*

Institute of Exercise Training and Sport Informatics, German Sport University, Cologne, Germany

ARTICLE INFO

Keywords:
Virtual reality
Training load
External
Internal
Seniors
Step training

ABSTRACT

Objective: Exercise-based fall prevention programs mainly refer to multimodal and challenging balance exercises. Individual load monitoring and interpretations are crucial to enable adequate adaptation responses on the individual level. Thus, assessing internal responses to external stimuli throughout an intervention period need to be adequately addressed. The aim of this secondary analysis of a 3-armed randomized controlled trial was to analyze internal and external loads of unstable vs. stable balance Exergame training in healthy seniors. We intended to elucidate whether differences of external and internal load criteria occur over the intervention period.

Methods: A total of 51 healthy seniors (females: n = 34; males: n = 17; age: 69 ± 6 years; BMI: 27 ± 5) were allocated to either volitional stepping (VOL), volitional stepping under unstable conditions (VOL + US) or an inactive control group (CON). VOL and VOL + US completed 8 weeks of Exergame based step training (three weekly sessions, 45 min each) using the Dividat Senso device. Twelve different balance Exergames were used, consisting of virtual reality like video games. The original nonswinging, stable platform was employed for VOL, whereas VOL + US used an adapted Senso mounted on a swinging Posturomed Rack. The instability level was increased for VOL + US only every second week. External (game scores) and internal (perceived efforts, using the rated perceived exertion scale (RPE)) load measures were individually recorded for every session. Statistical analysis was carried out using linear mixed-effects modelling.

Results: Although VOL + US completed similar games at identical training volumes under unstable conditions, the achieved game scores did not significantly differ between both training groups ($p = 0.71$). Both intervention groups notably improved their game scores over the 8 training weeks ($p < 0.01$). A significant time x group interaction effect was observed for perceived effort ($p < 0.01$), serving as an internal load measure. Subsequent post-hoc testing revealed significant greater perceived exertion values in each of the first 7 weeks ($p < 0.05$) in VOL + US compared to VOL. No between-group differences were found for RPE in week 8. Whereas RPE values in VOL + US decreased over time (week 1: 4.6 ± 1.9; week 8: 3.1 ± 1.6), VOL indicated similar RPE values for all weeks (week 1: 3.1 ± 1.3; week 8: 2.9 ± 1.4). A detailed analysis of all twelve games revealed that differences in perceived exertion depend on the game content: in 75% of the involved games the RPE level was significantly higher in VOL + US compared to VOL ($p < 0.05$).

Conclusion: Monitoring internal and external loads on individual level are paramount for gaining adequate training adaptations. Our results indicate that between-group differences in perceived efforts a) can funnel over time, b) depend on game content and c) do not necessarily affect overall scoring. Future studies should individually employ and monitor measures of perceived efforts to guarantee an adequate challenge to the balance system within exercise-based fall prevention programs.

1. Introduction

Demographic changes will lead to increasing proportions of older adults in western societies until the end of the 21st century (Lutz et al., 2008). These changes will have remarkable impact on direct and indirect health care costs (Stevens et al., 2006). Physical inactivity (PI)

* Corresponding author at: German Sport University Cologne, Institute of Exercise Science and Sport Informatics, Am Sportpark Muengersdorf 6, Cologne 50933, Germany.

E-mail address: r.rein@dshs-koeln.de (R. Rein).

https://doi.org/10.1016/j.exger.2020.111037
Received 10 March 2020; Received in revised form 17 July 2020; Accepted 24 July 2020
Available online 28 July 2020

serves as an independent risk factor for the development of numerous non-communicable diseases and can be seen as a driving factor of adverse disease-related developments on individual and public health care level (Blair, 2009). In turn, adequate physical activity contributes to physical, mental, and social well-being. Beside stress management and proper nutrition, strength, endurance, balance and flexibility have been mainly proposed to successfully maintain a healthy, fit and independent lifestyle (Nelson et al., 2007). These recommendations sum up to a cumulative weekly exercise training volume of more than 500 min. As a consequence, efficient, less time-consuming more integrative (Donath et al., 2016a) as well as multimodal (Sherrington et al., 2019) training frameworks have been suggested in recent years. Along with a challenging and specific training stimulus (Giboin et al., 2015), successful exercise training programs need to be tailored to individual circumstances and preferences, intended adaptations, and potential barriers (Fischer et al., 2019; Hecksteden et al., 2018).

Exercise training interventions mainly focus on improving physical performance and function. Adequate modifications of the applied training load and intensity are key prerequisites to sustain training progress with proper adaptation. Thereby, it seems crucial to monitor whether and how participants respond to a single and cumulative training stimuli compared to control conditions. Most training studies are designed as two- or three-armed trials, reflecting the gold standard of evidence-based decision making (Hecksteden et al., 2018). According to the consort requirements (Schulz et al., 2010), reporting of these interventions should provide adequate information on training content and overload criteria (volume, intensity, type, frequency). Often, these measures are reported only approximately, as there is often no close monitoring during each individual training session. However, continuous monitoring and assessment of training volume and intensity are essential to understand training-performance relationships and properly evaluate training responses (Mujika, 2017). In this regard, the integrative evaluation of internal and external loads is paramount for successful and efficient exercise training programming.

“External load monitoring” assesses the performed work using various tools and devices, such as accelerometers, time-motion analysis, and GPS (Impellizzeri et al., 2019; Bourdon et al., 2017). “Internal load monitoring” mirror individual physiological reactions of the body during a specific training session (Impellizzeri et al., 2019). Questionnaires, perceived exertion scales (Borg, 1982), biosensors or heart rates are well known methods employed to record internal loads (Halson, 2014). These measures enable both immediate training feedback and training adjustments to ensure intended training adaptations.

In the field of neuromuscular training, training monitoring is considered more challenging as, for example, balance training elicits various spinal and supra-spinal responses (Taube et al., 2008). Accordingly, these responses are difficult to monitor with feasibly applicable tools and devices. Available evidence suggests that Exergames in general and stepping interventions in particular improve a variety of fall-related risk factors and functional mobility outcomes (Donath et al., 2016b; Okubo et al., 2017; Swanenburg et al., 2018). The American College of Sports Medicine (ACSM) defined Exergaming as technology-driven physical activities using a virtual reality environment, like video game plays (Witherspoon, 2013). Exergaming seems to be a suitable interface to apply appealing and effective exercises and enable the assessment of individual training response data to adapt training stimuli to ensure sufficiently challenging training environments.

Against this background, the aim of the present study was to monitor, collect and compute each training task and session of volitional, stable stepping training compared to volitional, unstable stepping training. We aimed at elucidating whether differences of external load (game scores) and internal load (perceived efforts) criteria occur over the time course of the intervention. The obtained information may help to improve training programming and adjustment in order to provide intended, specific and challenging training stimuli.

2. Methods

2.1. Study design

The initial study (Morat et al., 2019) was designed as a longitudinal randomized controlled trial with three parallel groups. A total of 51 healthy seniors were allocated to either volitional stable stepping (VOL), volitional stepping under unstable conditions (VOL + US) or an inactive control group (CON) using the minimization method (strata: sex, age, BMI, 6-min walk, dynamic balance performance). The treatment groups underwent 8 weeks of Exergame based step training either on the original rigid platform (VOL) or mounted on an unstable Posturomed Rack (VOL + US). Pre- and Post-testing during the initial study included static, reactive, functional, dynamic balance and mobility assessment. All participants received all relevant study information prior to the study start and signed a written informed consent. The study complied with the Declaration of Helsinki and was further approved by the local ethics committee at the German Sport University Cologne (approval number 134/2018). The study protocol was not registered prior to the start of the study in a clinical trial registry.

2.2. Participants

Initially, 62 community-dwelling seniors were recruited by advertisements placed in local newspapers. Those showing any acute cardiovascular, psychological, neurological or orthopedic (also symptomatic knee- or hip-prostheses are not allowed) diseases were excluded. Participants had to be retired and older than 60 years. Finally 51 seniors (69.4 ± 5.6 years) were enrolled in the study (CON: n = 17; VOL: n = 17; VOL + US: n = 17). CON was not considered for this analysis, so the data set contains 34 participants. The Anthropometric baseline characteristics did not significantly differ between the three groups ($p > 0.05$)(See Table 1).

2.3. Stepping intervention

Participants of the intervention groups underwent Exergame based step training in small groups of three persons. The training lasted 8 weeks and was supervised three times a week (40 min each) by two qualified research assistants. Both intervention groups trained on the Dividat Senso device (Senso, Dividat, Schindellegi, Switzerland). The Dividat Senso is a training platform (1.13 m∗1.13 m) with force sensors linked to a screen (Fig. 1). Position and timing information are detected by electronic sensors embedded into the platform which provides the participants with real- time visual and auditory feedback concerning their performance. The platform sensors capture step forces along two dimensions. VOL performed the step training under stable conditions using the originally rigid platform of the Dividat Senso System. In contrast, VOL + US trained on a combined stepping device of the Dividat Senso system with the unstable swinging Posturomed Rack (Haider Bioswing, Pullenreuth, Germany). The Posturomed Rack is suspended on eight steel springs, allowing the platform to swing freely along the horizontal plane. Fixating or loosening of the springs in each

Table 1
Baseline characteristics given as means (standard deviations). BMI = body mass index, Physical activity assessed by the Freiburger Physical Activity Questionnaire (Frey et al., 1999).

Characteristics	VOL + US (n = 17)	VOL (n = 17)	CON (n = 17)	p-Value
Gender, female/male [n]	5/12	6/11	6/11	p = 0.920
Age [years]	68.2 (6.6)	69.8 (6.4)	71.4 (5.3)	p = 0.311
BMI [kg/m^2]	27.2 (4.8)	29.2 (5.7)	26.3 (3.4)	p = 0.197
Physical activity [h/week]	10.6 (5.4)	7.2 (4.7)	8.5 (4.6)	p = 0.183
Falls [n, past 12 months]	3	2	2	p = 0.853

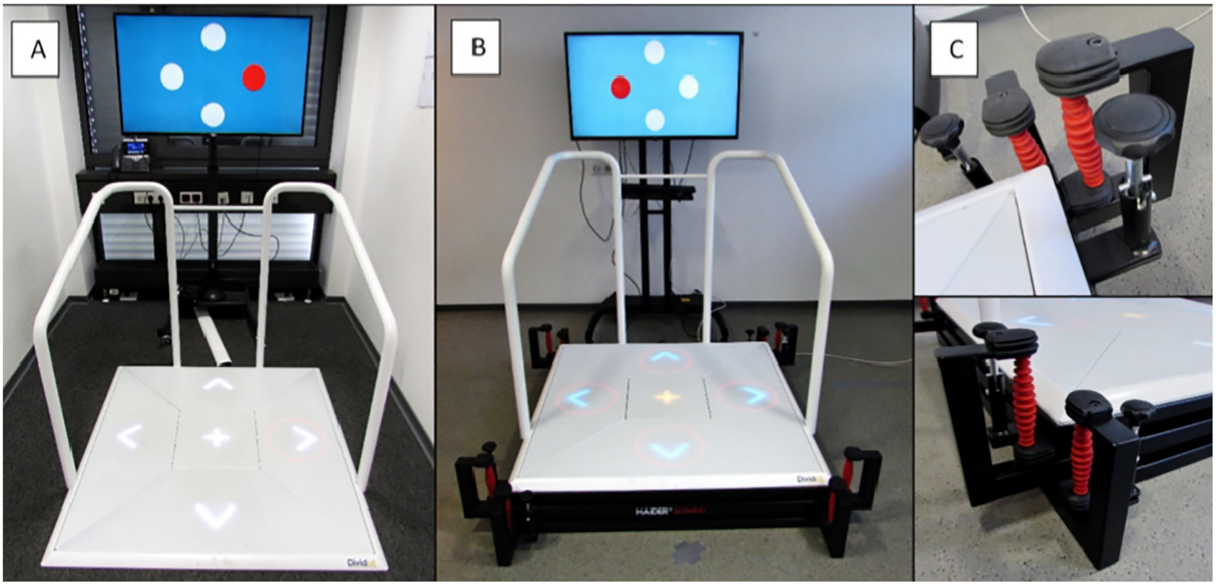


Fig. 1. A. Original Dividat Senso; B. Dividat Senso Swing (mounted on a Posturomed Rack): C. Close-up of the Posturomed Rack.

corner of the platform enables the adjustment of the degrees of freedom. Every second week, one additional spring was loosened by the supervisors for a continuous progression of instability. VOL + US participants started with three fixed springs and ended with four opened springs.

In total, 12 different stepping Exergames were played during the eight-week intervention. Games were selected to provide varying training stimuli targeting both cognitive and motor capacity. Stepping movements was the main motor task in nine of eleven games ("Birds", "Divided", "Flexi", "Habitats", "Hexagon", "Simon", "Simple", "Snake", "Targets" and "Tetris"). All Exergames were designed to train the following cognitive abilities: cognitive flexibility, divided and selective attention, mental rotation, postural control and visuospatial working memory. The main motor task in the other two games ("Rocket" and "Ski") was shifting body weight and endurance with no additional cognitive task. The participants were expected to play the games in a standing position without any support. The arms were akimbo. One exercise session lasted 40 min, with a net gaming time of 10–12 min per person. For each session, three shorter games (30–75 s.) or two games with one longer duration (150 s.) were preselected. As the participants alternated after each round, there was a standardized break between each game of 2–3 min, during which the participants either sat or stood next to each other. The same game selection procedure was used across both groups. The difficulty of instability increased only by opening another spring in the VOL + US group. Some of the games automatically adjusted game difficulty based on individual error rates through increased game velocity. However, game difficulty did not automatically increase in-between training sessions. At the end of each game, participants were asked to report their perceived exertion level using the Borg CR-10 scale, serving as measure of internal load. In general, there was no performance-based adaptation algorithm of the Dividat device. The progression was based on the perceived effort made by the participants and was manually adapted afterwards by the research assistants.

2.4. Game scores/assessment

The game scores, point rate, and reaction time of the exercise sessions were recorded for each game and individual in training logs and represent the focus of the present analysis. In all sessions, each individual attempt was documented as a part of an individual training response log for each participant. For statistical analysis, the scores for each game were individually converted into z-scores to allow comparison across games.

2.5. Testing procedure

All baseline and post-intervention measurements were conducted in the primary study (Morat et al., 2019) and serve as a basis for this analysis. The same research assistants performed all measurements at a comparable time of day while participants received one-on-one attention. Assessor blinding was not possible due to limited personnel. The participants were aware of the training mode but did not try or see the training of the other group before the end of the study. The measurements included balance, mobility and strength testing. Reactive balance performance was assessed as postural sway upon perturbation on the Posturomed and functional balance performance was tested for both legs with the Y-balance test (Y balance test kit, Perform better, Graefelfing, Germany). The Timed Up and Go Test (TUG) combines the measure of functional balance and mobility and was performed by walking instead of running as fast as possible (Podsiadlo and Richardson, 1991). Additionally, the participants performed the TUG in a cognitive dual-task condition by counting backwards from a number between 20 and 100 in steps of three, and in a motor dual-task condition carrying a cup of water as additional task. Maximal strength and the maximal rate of force development were measured in isometric conditions in a leg curl and a leg extension machine (Edition-Line, gym80, Gelsenkirchen, Germany). For the calf strength, participants performed a heel rise test for 30 s.

This is a secondary analysis of a three-armed randomized controlled trial. For detailed information of the testing procedures and results please compare "Effects of stepping Exergames under stable versus unstable conditions on balance and strength in healthy community-dwelling older adults: A three- armed randomized controlled trial" published in Experimental Gerontology (2019) by Morat and colleagues.

2.6. Statistical analysis

All statistical analyses were carried out using the software package (R Core Team, 2017). Statistical effects were analyzed using a linear mixed-effects modelling approach. Participants were modelled as random effects to account for repeated measures effects whereas group (VOL, VOL + US) and week served as fixed effects. The group effect

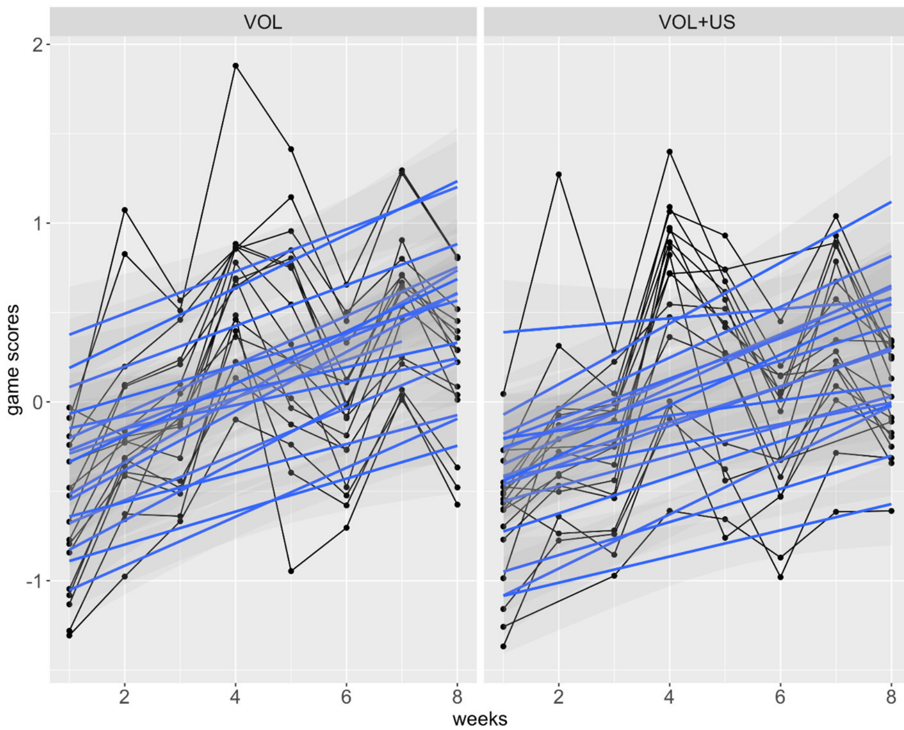


Fig. 2. Averaged z-scaled game scores for each individual for VOL and VOL + US groups over the 8 weeks. Black lines = individual mean, blue lines = linear trend for each participant. VOL = volitional stepping under stable conditions, VOL + US = volitional stepping under unstable conditions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

was coded using effects dummy coding (Bates et al., 2015). Between-groups differences were tested by type III Wald-likelihood ratio tests using x^2 approximation (Fox and Weisberg, 2019). In case of p-values for main- or interaction effects falling below the alpha level of $p < 0.05$, post-hoc testing was performed using expected marginal means (Searle et al., 1980; Lenth, 2019). Effect sizes were additionally calculated following Westfall et al. (2014) and Brysbaert and Stevens (2018). The significance level for all tests was set at the alpha level of 0.05. All outcome measures are given as means with standard deviations (± SD).

3. Results

3.1. Game scores

The transformed game scores for each participant over the 8 weeks are displayed in Fig. 2. VOL + US completed the same games at identical training volume under instable conditions and played the games with four loosened springs during the final two weeks. Statistical testing indicated no significant differences between the game scores of the two groups ($x^2(1) = 0.14$; $p = 0.71$; $d = 0.13$). All participants in both groups increased their game scores over the course of the 8 weeks ($x^2(1) = 208.37$; $p < 0.01$) whereas the time x group interaction effect was not significant $x^2(1) = 2.90$; $p = 0.09$. Thus, increases in game score performance were similar across groups despite the more demanding postural requirements for the VOL + US group.

3.2. Perceived effort (RPE)

Fig. 3 shows the perceived exertion level of each participant during all games over the treatment period. For VOL, the average level remained relatively stable whereas the VOL + US group depicts much larger variation during the treatment duration. Statistical analysis of the RPE scores revealed no significant main effect ($p = 0.74$) between both intervention groups and across all Exergames and weeks. The mean RPE score was 4.1 ± 1.9 in VOL + US and 3.1 ± 1.3 in VOL. A significant time x group interaction effect was found (x^2 (7) = 176; $p < 0.01$). Subsequent, post-hoc testing revealed greater perceived exertion values for each of the first seven weeks in VOL + US compared to VOL (week 1: $p < 0.01$, $d = 0.76$; week 2: $p < 0.01$, $d = 0.72$; week 3: $p < 0.01$, $d = 0.91$; week 4: $p = 0.02$, $d = 0.29$; week 5: $p < 0.01$, $d = 0.67$; week 6: $p = 0.02$, $d = 0.45$; week 7: $p < 0.01$, $d = 0.52$). As the exertion values in VOL + US decreased significant over time (week 1: 4.6 ± 1.9; week 8: 3.1 ± 1.6), VOL indicated similar RPE values over time without major changes (week 1: 3.1 ± 1.3; week 8: 2.9 ± 1.4) and an insignificant difference between the both groups in week 8 was observed ($p = 0.74$, $d = 0.04$).

Table 2 shows a detailed analysis of all 12 Exergames with perceived exertion values over the 8 weeks and indicates that the perceived exertion differed depending on the game content. Statistical testing indicated a significant time x group interaction effect $x^2(11) = 367.165$; $p < 0.001$. VOL + US showed greater RPE levels in most games (8 games = 75%) compared to VOL. Post-hoc testing indicated significant differences between the following eight games: “Birds” ($p = 0.02$); “Flexi” ($p < 0.01$); “Hexagon” ($p < 0.01$); “Simon” ($p < 0.01$); “Ski” ($p < 0.01$); “Snake” ($p < 0.01$); “Targets” ($p < 0.01$); “Tetris” ($p < 0.01$). In the other four Exergames “Divided” ($p = 0.74$); “Habitats” ($p = 0.19$); “Rocket” ($p = 0.33$) and “Simple” ($p = 0.85$) the RPE values are similar between VOL + US and VOL and insignificant (Table 2).

4. Discussion

To the best of the authors' knowledge this is the first study investigating the association between internal and external training load in two different neuromuscular training programs for older adults. The main aim of this analysis was to monitor, collect and compute each training task and session of both groups and to elucidate whether differences between the external and internal load occur over the time of intervention. Increases in game score performances, as external load, were similar across groups and no significant differences have occurred. Perceived exertion level, as internal load, remained relatively stable for VOL, whereas VOL + US revealed significant higher perceived exertion values in each of the first seven weeks compared to VOL. In the last week of the intervention the internal load level was similar in both groups and no significant difference was observed. It seems that

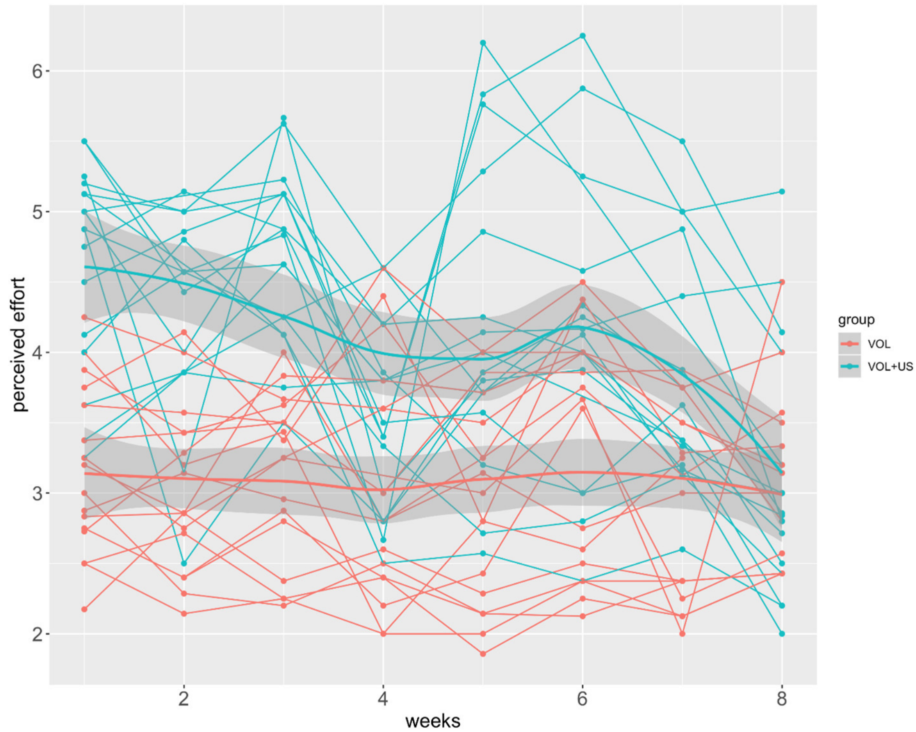


Fig. 3. Individual perceived exertion values for both groups over the 8 weeks. Data are presented as means for each individual. Thin lines: individual participants. Thick lines: Local polynomial smoother (degree = 2, span = 0.75).

VOL + US adapted to the highly challenging balance training over the eight weeks. The results of the primary study on reactive balance underlines these findings of the RPE development in VOL + US. Post-hoc testing revealed a statistically significant improvement ($p = 0.016$)) over the eight weeks of VOL + US for total postural sway upon perturbation in contrast to the VOL group, although the effect size was rather small (Standardized Mean Difference = 0.3). Nevertheless, VOL + US improved their reactive balance over the treatment period and this improved performance might have offset the increased difficulty of the task yielding the decrease of their internal load through perceived exertion. Thus, future studies and unstable training platforms should incorporate monitoring procedures to track training progression and variation to allow inter-individual progressive modulation of internal load conditions.

The present results support the argument put forward by Hecksteden et al. (2018) that both parameters should be taken into account simultaneously when designing and monitoring individualized training programs. The concept of "Training load" is defined as the combination of external (what a person could actually do) and internal training load (acute individual experience of training) (Foster et al.,

Table 2
Changes of perceived exertion from first to last time playing each Exergame. Data are presented as means with standard deviations.

Game	Group	First mean ± SD	Last mean ± SD	Between-group comparison	Effect size
Birds	VOL	2.00 ± 0.00	1.81 ± 0.39	**p = 0.0173**	**d = 0.4**
	VOL + US	2.64 ± 0.73	2.33 ± 0.80		
Divided	VOL	2.13 ± 0.34	2.23 ± 0.43	p = 0.7372	d = 0.05
	VOL + US	2.00 ± 1.10	2.19 ± 0.73		
Flexi	VOL	3.24 ± 0.95	3.07 ± 0.71	**p < 0.0001**	**d = 0.98**
	VOL + US	5.00 ± 1.48	3.79 ± 0.95		
Habitats	VOL	2.40 ± 0.62	2.09 ± 0.52	p = .1936	d = 0.21
	VOL + US	3.73 ± 0.78	2.00 ± 0.76		
Hexagon	VOL	3.46 ± 1.62	2.56 ± 1.07	**p < .0001**	**d = 1.2**
	VOL + US	5.56 + 2.17	3.73 ± 2.52		
Rocket	VOL	3.40 ± 1.28	5.08 ± 1.75	p = .3338	d = 0.18
	VOL + US	4.31 ± 2.08	3.55 ± 2.22		
Simon	VOL	4.30 ± 1.16	3.75 ± 1.45	**p < .0001**	**d = 1.38**
	VOL + US	5.06 ± 0.43	5.14 ± 1.14		
Simple	VOL	1.67 ± 0.60	1.64 ± 0.48	p = .8468	d = 0.04
	VOL + US	1.94 ± 0.24	1.53 ± 0.51		
Ski	VOL	2.92 ± 0.74	2.55 ± 1.09	**p < .0001**	**d = 0.71**
	VOL + US	4.94 ± 1.77	2.93 ± 1.45		
Snake	VOL	3.80 ± 1.24	3.08 ± 0.84	**p < .0001**	**d = 0.82**
	VOL + US	4.94 ± 0.66	4.38 ± 1.78		
Targets	VOL	2.20 ± 0.55	2.31 ± 0.83	**p < .0001**	**d = 0.72**
	VOL + US	3.53 ± 0.63	3.06 ± 1.31		
Tetris	VOL	3.77 ± 1.20	4.13 ± 1.42	**p < .0001**	**d = 1.54**
	VOL + US	6.81 ± 2.24	6.53 ± 1.65		

2017). Various monitoring methods for measuring "training load" have been used across sports disciplines (RPEs). These methods range from feasible measurements of heart rate and training volume to self-assessment of training sessions by perceived exertion for example (Foster et al., 2017). Recording internal load through perceived exertion has the advantage to quantify load irrespective of mode or location. Although RPE is a rather subjective tool, previous studies have proven its validity as a means of regulating exercise intensity (Eston, 2012; Haddad et al., 2017). In the present study, the different Exergames targeted different neuromotor abilities and thus the recorded RPE values were influenced by both cognitive and motor requirements. Exergame technology provides a new balance training for the elderly by the combination of cognitive and motor challenges of dynamic postural control within a complex multimodal task (Van Diest et al., 2013). In addition, VOL + US had more demanding postural requirements due to increasing instability level, by what the variation of the perceived effort was much bigger than in VOL during the time of intervention. This can be seen in the results of the first seven weeks with significant greater perceived exertion values in VOL + US. The significant decrease in exertion values over time in VOL + US, as well as the insignificant difference between both intervention groups in the last week (week 8), combined with similar increases in game scores proves that RPE was a sensitive tool for internal load monitoring. It seems that the internal load monitoring by the RPE scale is in this case a successful and useful tool for the right individual adaptation for performance improvements. At that time the study was conducted, there was the opportunity to set up a profile for every participant. Due to lack of time and resources, we did not consider this, moreover this function did not bring any added value at that time. Dividat AG updated their system with a new actual version with the opportunity to an individual automatic training load adjustment in real time, performance reports over periods and a rich source of data for scientists. The participants will receive an individual feedback directly after each game. This is a huge improvement especially for monitoring training sessions, progression and load adjustment for an individual challenging training stimuli and should be investigated in further studies.

The intervention period was rather short with 8 weeks and should be longer in further research. The duration of exercise training interventions as well as the overall training volume are important factors influencing training adaptations and transfer effects (Hecksteden et al., 2018). The typical duration of exercise studies ranges between 4 and 12 weeks. Lesinski et al. (2015) revealed that a training period of 11–12 weeks induced the greatest effects on both overall balance performance as well as more specific measures of static steady-state balance. Transfer effects in balance studies should be treated with caution (Giboin et al., 2015; Kümmel et al., 2016; Donath et al., 2017) as for example a meta-analysis by Kümmel et al. (2016) revealed strong specificity of balance training adaptations. Accordingly, balance training improves performance only in the explicit trained balance tasks and the transfer of performance gains from trained to untrained balance tasks is either small or negligible (Kümmel et al., 2016). The various cognitive and motor tasks underlying the 12 games used in the present study covered a wide spectrum of balance tasks targeting fall prevention. Thus, this provides a possible explanation of the transfer effects found in strength and reactive balance in the initial study (Morat et al., 2019).

The most important variable for training adaptations seems to be the overall training volume and the assumption that greater total training volumes lead to greater training effects seems rather obvious. In this regard, Lesinski et al. (2015) proposed that a total number of 36–40 training sessions over several months usually lead to strong exercise training effects and is most effective in overall balance performance. The 8 weeks training period with a total of 24 sessions per participant used in the present study therefore may not be considered optimal. However, the results provide a first point of departure towards an assessment of "training load" in the domain of neuromuscular training intervention programs. The majority of training studies in the literature report numerous training parameters (e.g. frequency, intensity, time, type, volume) of the adopted intervention (Hecksteden et al., 2018; Schulz et al., 2010). But in most cases these parameters are not monitored in detail and often lacks information about individual training responses. Accordingly, the adequacy of individual training responses is neither reported nor can be assessed and justified. In this respect, systems based on VRT with Exergames facilitate the recording and documentation of all individual trials for each participant (compare also Swanenburg et al., 2018). Exergames enable detailed monitoring of training progression which is currently an underutilized aspect in exercise training interventions (Hecksteden et al., 2018). Increasingly, studies in motor learning also adapt trial-based approaches to map out intra-individual and inter-individual changes during motor learning (Chow et al., 2008). In fact, studies focusing specifically on the effect of load progression in exercise training studies are rare. Relating to the challenging balance training for fall prevention in the elderly (Sherrington et al., 2019), the VOL + US group was exposed to a continuous progression of instability over the 8 weeks. Despite the increasing instability, the results demonstrate improvements in the game scores as well as a decrease of the perceived effort over the intervention duration. The lack of group differences may result from the high between-group variability. Although there was some inter-individual variability, all participants in both groups increased their game score over the training period. Thus, future studies should consider reporting trial-based measures to increase opportunities for individualized modelling of training responses.

A potentially interesting framework in this regard might be the challenge point hypothesis proposed by Guadagnoli and Lee (2004). The challenge point hypothesis poses that individual learning rates are dependent on the interplay between individual skill level and task difficulty following a U-relationship. When the task is either too easy or too difficult learning rates are decreased. Therefore, to support optimal learning rates task difficulty should continually adapt throughout treatment phases. Guadagnoli and Lee (2004) frame task difficulty with respect to the amount of information available to the learner to regulate behavior. Accordingly, there is an optimal amount of potential interpretable information for each individual hence the optimal challenge points. With increasing skill over time, the performance expectations (greater game scores) increase. Thus, to expose learners to new learning challenges, increased information should be available which can be regulated through task difficulty (e.g. in the present example through additional number of opened springs) (compare also Onla-Or and Winstein, 2008). The optimal challenge point depends on the skill level of each individual learner and accordingly the same external task difficulty results in different information amount across individuals (Guadagnoli and Lee, 2004). Ozgur et al. (2020) showed that older adults improved their game performance and demonstrated learning in various game configurations over time. Besides an increasing challenge level, feedback is an important factor to find adequate solutions for different motor problems and may positively influence motor learning by enhancing motivation and engagement (Levin, 2020). However, without monitoring internal and external load simultaneously, it is impossible to ensure an appropriate and individualized, optimal dosed training. Thus, to ensure actors are optimally challenged both loads should be monitored to maximize training effects.

Although the study allows first assumptions and insights about load monitoring, there are some limitations that have to be addressed. The short intervention period and the low training volume with in total 24 sessions were already mentioned, as well as the relatively small sample size. Relating to the original paper of Morat et al. (2019) future research should allocate subjects to the different study arms after completing all baseline measures. Despite using the minimization method for group allocation, there have been some baseline differences concerning balance performance. Only the following six strata criteria: sex, age, BMI, 6-min walk and dynamic balance performance were included.

VOL + US had the lowest reactive balance performance at baseline. This could have influenced and might have contributed to significant higher perceived exertion levels in the first weeks in VOL + US and significant improvements over time, despite the progression of instability. With regard to the newest version of the Dividat Senso System, researcher would have much more information and data about game scores, training load adjustment, performance improvement and the individual progression of every subject. Moreover, different levels of the games for an adequate challenge as used in the study of Litz et al. (2020) would have enriched the current analysis. These points should be considered in further research.

5. Conclusion

A key issue in exercise training interventions is that identical workloads can lead to different individual training responses and assessment. The present analysis underpins the relevance of internal and external load monitoring for successful individual exercise training programming. The participants should be neither over- nor under-challenged by the different games and instability levels. An individual training control adjustment is essential for the success of the specific training and possible transfer effects. Measures of internal load derived from perceived exertion and measures of external load in form of game scores are suitable methods in the field of Exergame based balance training and may serve as a proper tool for individual assessment. Additional advantages of Exergames are high adherence, the motivation of the participants and that their focus of attention is not on the movement itself, but on the pleasure and outcome of the movements in the game. Stepping Exergames, especially under unstable conditions with an increasing instability, offer a high specificity that might offer necessary motor skills to recover balance from trips and slips and promote healthy aging. These results further demonstrate how VRT enables individual and progressively challenging neuromuscular exercise training. For this purpose, perceived efforts should be gradually adapted in order to provide adequate training stimuli on an individual level.

Statement of ethics

The authors have no ethical conflicts to disclose.

Disclosure statement

The authors have no conflicts of interest to declare.

Funding sources

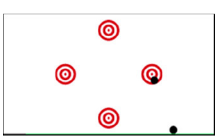
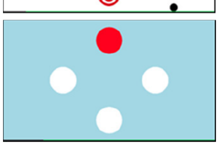
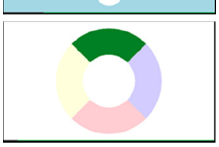
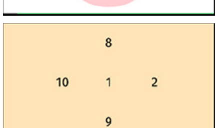
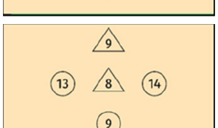
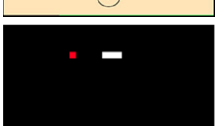
This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

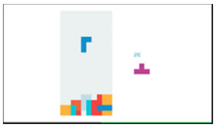
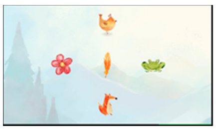
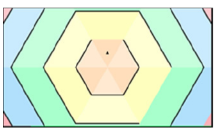
Author contribution statement

Lars Donath, Robert Rein and Julia Bakker conceived of the presented idea. The background information in the introduction section is from Lars Donath and Julia Bakker. Robert Rein designed the model and the computational framework and analyzed the data. The data visualization was carried out by Robert Rein and Julia Bakker. All authors discussed the results and contributed to the final manuscript.

The authors declare that they have no conflicts of interest. No specific grants from funding agencies in the public, commercial or not-for-profit sectors have been received. The local institutional Human Research Ethics Board granted ethical approval and the study followed the principles set out in the Helsinki declaration (2008). This manuscript is original and not previously published, nor is it being considered elsewhere until a decision is made as to its acceptability by the Experimental Gerontology Editorial Review Board. The publication has been approved by all authors.

Appendix A. Overview over the Dividat Senso games.

Game	Screenshot	Description
Targets		Four targets are displayed and black circles "fly" randomly from all directions across the screen. When the black circle hits the center of a target, the participant has to step in the respective direction of the target. The "Bullseye" (perfect match of the center) results in highest scoring.
Divided		Four white dots are displayed on the screen. As soon as a dot randomly turns red, the participant has to quickly step in the respective direction. In between the appearance of red dots, a high or a low acoustic signal is sometimes presented, requiring a step forward (high) or backward (low).
Simon		In the memorization phase, a sequence of acoustic signals is played. The respective movement directions are simultaneously indicated by lighting up the corresponding color in the circle. In the subsequent recall phase, the participant has to reproduce the sequence by stepping in the respective directions in the right order. The sequence starts with one signal and after each successful recall, the sequence adds one more signal.
Flexi		a) A number is displayed in the center of the screen, surrounded by four additional numbers. Starting from the number in the center, the participant steps in the direction of the number that is the next in line. In the top example, stepping to the right "2" would be correct. b) In a higher level of the game, the numbers are additionally framed by different shapes (e.g. triangles, circles). Starting from the number in the center, the participant steps in the direction of the number that is next in line and is surrounded by a different shape. In the lower example, stepping backwards to the "9" in a circle would be correct.
Snake		A white snake winds its way over the screen and is supposed to "eat" the red squares occurring at random places on the screen. The snake is navigated by the participant by stepping to the required direction. With each "eaten" square, the snake becomes longer. The snake is not allowed to touch its own tail.

Tetris		Different shapes “fall” from the top of the screen. The participant has to rotate the shapes by 90° per step (forward) and to shift them (step right or left). All segments need to be placed without or with a minimum of gaps at the lower edge of the screen. As soon as a row is complete, it disappears and the top rows move down. The game is over if a row touches the top of the screen.
Habitats		Four habitats are presented on the screen. An animal appears in one of the habitats. The participant has to step in the respective direction only if the animal fits the habitat. In this example, a fish appears in the upper “sky” habitat, where it does not fit. Thus, the participant is not required to step.
Birds		Different items are displayed - one in the center of the screen and four surrounding it. One of the four items always matches the one presented in the center. The participant is supposed to step in the respective direction. In this example, the “feather” belongs to the bird, so stepping forward would be correct.
Hexagon		Hexagons in increasing sizes are displayed on the screen. In the inner of the center hexagon, a small black arrow is shown. The hexagon shapes move down while the arrow stays at its place. The participant has to turn the hexagons left or right by stepping in the respective direction, so that the open side of the hexagon is at the top and the arrow does not touch a wall of a hexagon, while it moves down.

References

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models Usinglme4. J. Stat. Softw. 67 (1). https://doi.org/10.18637/jss.v067.i01.

Blair, S.N., 2009. Physical inactivity: the biggest public health problem of the 21st century. Br. J. Sports Med. 43 (1), 1–2.

Borg, G.A., 1982. Psychophysical bases of perceived exertion. Med. Sci. Sports Exerc. 14 (5), 377–381.

Bourdon, P.C., Cardinale, M., Murray, A., Gastin, P., Kellmann, M., Varley, M.C., Cable, N.T., 2017. Monitoring athlete training loads: consensus statement. International Journal of Sports Physiology and Performance 12 (s2), S2–161-S2-170. https://doi.org/10.1123/ijspp.2017-0208.

Brysbaert, M., Stevens, M., 2018. Power analysis and effect size in mixed effects models: a tutorial. J. Cogn. 1 (1).

Chow, J.Y., Davids, K., Button, C., Rein, R., 2008. Dynamics of movement patterning in learning a discrete multiarticular action. Mot. Control. 12 (3), 219–240. https://doi.org/10.1123/mcj.12.3.219.

Donath, L., van Dieën, J., Faude, O., 2016a. Exercise-based fall prevention in the elderly: what about agility? Sports Med. 46 (2), 143–149. https://doi.org/10.1007/s40279-015-0389-5.

Donath, L., Rössler, R., Faude, O., 2016b. Effects of virtual reality training (Exergaming) compared to alternative exercise training and passive control on standing balance and functional mobility in healthy community-dwelling seniors: a meta-analytical review. Sports Med. 46 (9), 1293–1309. https://doi.org/10.1007/s40279-016-0485-1.

Donath, L., Roth, R., Zahner, L., Faude, O., 2017. Slackline training (balancing over narrow nylon ribbons) and balance performance: a meta-analytical review. Sports Med. 47 (6), 1075–1086. https://doi.org/10.1007/s40279-016-0631-9.

Eston, R., 2012. Use of ratings of perceived exertion in sports. International Journal of Sports Physiology and Performance 7 (2), 175–182. https://doi.org/10.1123/ijspp.7.2.175.

Fischer, X., Kreppke, J.-N., Zahner, L., Gerber, M., Faude, O., Donath, L., 2019. Telephone-based coaching and prompting for physical activity: short- and long-term findings of a randomized controlled trial (Movingcall). Int. J. Environ. Res. Public Health 16 (14), 2626. https://doi.org/10.3390/ijerph16142626.

Foster, C., Rodriguez-Marroyo, J.A., de Koning, J.J., 2017. Monitoring training loads: the past, the present, and the future. International Journal of Sports Physiology and Performance 12 (s2), S2–2-S2-8. https://doi.org/10.1123/ijspp.2016-0388.

Fox, J., Weisberg, S., 2019. An R Companion to Applied Regression (Third).

Frey, I., Berg, A., Grathwohl, D., Keul, J., 1999. Freiburg questionnaire of physical activity—development, evaluation and application. Sozial-und Praventivmedizin 44 (2), 55–64. https://doi.org/10.1007/bf01667127.

Giboin, L.-S., Gruber, M., Kramer, A., 2015. Task-specificity of balance training. Hum. Mov. Sci. 44, 22–31. https://doi.org/10.1016/j.humov.2015.08.012.

Guadagnoli, M.A., Lee, T.D., 2004. Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. J. Mot. Behav. 36 (2), 212–224. https://doi.org/10.3200/jmbr.36.2.212-224.

Haddad, M., Stylianides, G., Djaoui, L., Dellal, A., Chamari, K., 2017. Session-RPE method for training load monitoring: validity, ecological usefulness, and influencing factors. Front. Neurosci. 11, 612.

Halson, S.L., 2014. Monitoring training load to understand fatigue in athletes. Sports Med. 44 (S2), 139–147. https://doi.org/10.1007/s40279-014-0253-z.

Hecksteden, A., Faude, O., Meyer, T., Donath, L., 2018. How to construct, conduct and analyze an exercise training study? Front. Physiol. 9. https://doi.org/10.3389/fphys.2018.01007.

Impellizzeri, F.M., Marcora, S.M., Coutts, A.J., 2019. Internal and external training load: 15 years on. International Journal of Sports Physiology and Performance 14 (2), 270–273. https://doi.org/10.1123/ijspp.2018-0935.

Kümmel, J., Kramer, A., Giboin, L.-S., Gruber, M., 2016. Specificity of balance training in healthy individuals: a systematic review and meta-analysis. Sports Med. 46 (9), 1261–1271. https://doi.org/10.1007/s40279-016-0515-z.

Lenth, R., 2019. Emmeans: Estimated Marginal Means, Aka Least-Squares Means. R Package V. 1.3. 4.

Lesinski, M., Hortobagyi, T., Muehlbauer, T., Gollhofer, A., Granacher, U., 2015. Effects of balance training on balance performance in healthy older adults: a systematic review and meta-analysis. Sports Med. 45, 1721–1738. https://doi.org/10.1007/s40279-015-0375-y.

Levin, M.F., 2020. What is the potential of virtual reality for post-stroke sensorimotor rehabilitation? Expert. Rev. Neurother. 20 (3), 195–197. https://doi.org/10.1080/14737175.2020.1727741. 2020.

Litz, E., Ball, C., Jansen, C.P., Werner, C., De Bruin, E., Hauer, K., 2020. Validation of a motor-cognitive assessment for a stepping Exergame in older adults: use of game-specific, internal data stream. Games for Health Journal 9 (2), 95–107. https://doi.org/10.1089/g4h.2019.0081.

Lutz, W., Sanderson, W., Scherbov, S., 2008. The coming acceleration of global population ageing. Nature 451 (7179), 716–719. https://doi.org/10.1038/nature06516.

Morat, M., Bakker, J., Hammes, V., Morat, T., Giannouli, E., Zijlstra, W., Donath, L., 2019. Effects of stepping exergames under stable versus unstable conditions on balance and strength in healthy community-dwelling older adults: a three-armed randomized controlled trial. Exp. Gerontol. 127, 110719. https://doi.org/10.1016/j.exger.2019.110719.

Mujika, I., 2017. Quantification of training and competition loads in endurance sports: methods and applications. International Journal of Sports Physiology and Performance 12 (s2), S2–9-S2-17. https://doi.org/10.1123/ijspp.2016-0403.

Nelson, M.E., Rejeski, W.J., Blair, S.N., Duncan, P.W., Judge, J.O., King, A.C., Castaneda-Sceppa, C., 2007. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. Med. Sci. Sports Exerc. 39 (8), 1435–1445.

Okubo, Y., Schoene, D., Lord, S.R., 2017. Step training improves reaction time, gait and balance and reduces falls in older people: a systematic review and meta-analysis. Br. J. Sports Med. 51, 586–593. https://doi.org/10.1136/bjsports-2015-095452.

Onla-Or, S., Winstein, C.J., 2008. Determining the optimal challenge point for motor skill learning in adults with moderately severe Parkinson's disease. Neurorehabil. Neural Repair 22 (4), 385–395. https://doi.org/10.1177/1545968307313508.

Ozgur, A.G., Wessel, M.J., Olsen, J.K., Johal, W., Ozgur, A., Hummel, F.C., Dillenbourg, P., 2020. Gamified motor training with tangible robots in older adults: a feasibility study and comparison with the young. Front. Aging Neurosci. 12.

Podsiadlo, D., Richardson, S., 1991. The timed “up & go”: a test of basic functional mobility for frail elderly persons. J. Am. Geriatr. Soc. 39, 142–148. https://doi.org/10.1111/j.1532-5415.1991.tb01616.x.

R Core Team. R: A Language and Environment for Statistical Computing. 2017 [cited 04 July 2018]. In: R Foundation for Statistical Computing [Internet]. Vienna, Austria: R. Available from https://www.R-project.org/.

Schulz, K.F., Altman, D.G., Moher, D., 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Med. 8 (1). https://doi.org/10.1186/1741-7015-8-18.

Searle, S.R., Speed, F.M., Milliken, G.A., 1980. Population marginal means in the linear model: an alternative to least squares means. Am. Stat. 34 (4), 216–221. https://doi.org/10.1080/00031305.1980.10483031.

Sherrington, C., Fairhall, N.J., Wallbank, G.K., Tiedemann, A., Michaleff, Z.A., Howard, K., Clemson, L., Hopewell, S., Lamb, S.E., 2019. Exercise for preventing falls in older people living in the community. The Cochrane Database of Systematic Reviews 1, CD012424. https://doi.org/10.1002/14651858.CD012424.pub2.

Stevens, J.A., Corso, P.S., Finkelstein, E.A., Miller, T.R., 2006. The costs of fatal and non-fatal falls among older adults. Injury Prevention 12 (5), 290–295. https://doi.org/10.1136/ip.2005.011015.

Swanenburg, J., Wild, K., Straumann, D., Bruin, E.D.de., 2018. Exergaming in a moving

virtual world to train vestibular functions and gait; a proof-of-concept-study with older adults. Front. Physiol. 9. https://doi.org/10.3389/fphys.2018.00988.
Taube, W., Gruber, M., Gollhofer, A., 2008. Spinal and supraspinal adaptations associated with balance training and their functional relevance. Acta Physiol. 193 (2), 101–116. https://doi.org/10.1111/j.1748-1716.2008.01850.x.
Van Diest, M., Lamoth, C.J., Stegenga, J., Verkerke, G.J., Postema, K., 2013. Exergaming for balance training of elderly: state of the art and future developments. Journal of neuroengineering and rehabilitation 10 (1), 101.
Westfall, J., Kenny, D.A., Judd, C.M., 2014. Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. J. Exp. Psychol. Gen. 143 (5), 2020–2045. https://doi.org/10.1037/xge0000014.
Witherspoon, L., 2013. ACSM Information on Exergaming. American College of Sports Medicine.