Restorative Neurology and Neuroscience 29 (2011) 1–12
DOI 10.3233/RNN-2011-0599
IOS Press

Virtual reality based rehabilitation speeds up functional recovery of the upper extremities after stroke: a randomized controlled pilot study in the acute phase of stroke using the Rehabilitation Gaming System

Mónica da Silva Cameirão[a], Sergi Bermúdez i Badia[a,1], Esther Duarte[b] and Paul F.M.J. Verschure[a,c,*]

[a]*Laboratory of Synthetic Perceptive Emotive and Cognitive Systems (SPECS), Department of Technology, Universitat Pompeu Fabra, Barcelona, Spain*

[b]*Servei de Medicina Física i Rehabilitació, Hospitals del Mar i l'Esperança, Parc de Salut Mar. Sant Josep de la Muntanya 12, Barcelona, Spain*

[c]*Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain*

Abstract. *Purpose*: Given the incidence of stroke, the need has arisen to consider more self-managed rehabilitation approaches. A promising technology is Virtual Reality (VR). Thus far, however, it is not clear what the benefits of VR systems are when compared to conventional methods. Here we investigated the clinical impact of one such system, the Rehabilitation Gaming System (RGS), on the recovery time course of acute stroke. RGS combines concepts of action execution and observation with an automatic individualization of training.

Methods. Acute stroke patients ($n = 8$) used the RGS during 12 weeks in addition to conventional therapy. A control group ($n = 8$) performed a time matched alternative treatment, which consisted of intense occupational therapy or non-specific interactive games.

Results. At the end of the treatment, between-group comparisons showed that the RGS group displayed significantly improved performance in paretic arm speed that was matched by better performance in the arm subpart of the Fugl-Meyer Assessment Test and the Chedoke Arm and Hand Activity Inventory. In addition, the RGS group presented a significantly faster improvement over time for all the clinical scales during the treatment period.

Conclusions. Our results suggest that rehabilitation with the RGS facilitates the functional recovery of the upper extremities and that this system is therefore a promising tool for stroke neurorehabilitation.

Keywords: Acute stroke, rehabilitation, mirror neurons, virtual reality, Rehabilitation Gaming Station

[1] Current affiliation: Madeira Interactive Technologies Institute, Universidade da Madeira, Caminho da Penteada, 9020-105 Funchal, Madeira, Portugal.

*Corresponding author: Paul F.M.J. Verschure, SPECS, Department of Technology, Universitat Pompeu Fabra, Roc Boronat 138, 08018 Barcelona, Spain. Tel.: +34 935421366; Fax: +34 935422202; E-mail: paul.verschure@upf.edu.

1. Introduction

In recent years the use of technology based neurorehabilitation approaches has increased to face the high demands due to the increasing numbers of stroke victims (Mathers and Loncar 2006; Kalra 2009). One of these technologies is Virtual Reality (VR) that promises the development of effective rehabilitation environments as it provides rich controllable multi-modal simulation and the possibility for individualization. A number of studies showed evidence of the positive benefits of such systems in the rehabilitation of the paretic upper limb after stroke (Cameirao et al., 2008; Lucca, 2009). However, the impact of VR based approaches on recovery is not fully understood and its advantages with respect to traditional neurorehabilitation methods has not yet been convincingly proven (Lucca, 2009).

To address this issue, here we explored one specific VR based system, the Rehabilitation Gaming System (RGS). RGS is based on the assumption that task oriented action combined with the observation of virtual limbs that reflect the executed movements facilitates the functional reorganization of the neuronal systems directly or indirectly affected by stroke and functional recovery (see (Cameirao et al., 2010) for an extended description of the key assumptions behind RGS and their scientific grounding). This paradigm is based on the human Mirror Neuron System (MNS), a system that is active during both goal-oriented action execution and action observation performed with a biological effector (Rizzolatti and Craighero, 2004; Mukamel, Ekstrom et al., 2010; Rizzolatti and Fabbri-Destro, 2010). Hence, RGS proposes, based on the MNS literature, that a direct transduction channel exists between the perception of action and its execution and that this channel can be used to drive effective reorganization after stroke. Indeed, a similar, but not VR based approach has been very successful in the treatment of aphasia (Pulvermuller, 2005). Further, RGS is based on theoretical work that has elucidated the relatively uniform statistical learning mechanisms of the neocortex and the critical dependence of learning on the specific statistical structure of inputs (Olshausen and Field, 1996; Wyss, Konig et al., 2006). With respect to exploiting this feature of cortical learning this can be reformulated in terms of the specific and parametric control of the sensorimotor contingencies the brain is exposed to (O'Regan and Noe, 2001). The execution and observation of goal-oriented movements provides sensory feedback of one's actions in terms of movement patterns and movement outcomes, and such feedback has been shown to facilitate motor learning (Ungerleider et al., 2002; Krakauer, 2006). In addition, RGS exploits the phenomenon of behavioral feedback that proposes that the behavior dependent sampling of a sensory space optimizes learning (Verschure et al., 2003). Lastly, RGS is built around the notion of task dependent learning to exploit the role of neuromodulation in the regulation of plasticity (Sanchez-Montanes et al., 2000; Bao et al., 2001). Learning is tightly regulated by systems that relate to motivation and arousal [see for (Green and Bavelier, 2008) a review]. This implies that each user has to be encouraged to train at an optimal level of errors, avoiding boredom of only correct trials or the frustration of too many failures. For this reason RGS includes the, so called, Personalized Training Module (PTM), which adapts the task to the specific performance level of the user on a trial by trial basis to an average performance level of about 70% correct trials (Cameirao et al., 2010).

The version of RGS presented here engages the user with a game-like task, called Spheroids, that is based on the above principles. In Spheroids the user has to interact with upcoming spheres and perform specific movements from basic arm range movements, to grasping and object displacement, and release. Here we investigate the impact of RGS supported rehabilitation on the recovery time course after stroke in comparison to standard occupational therapy and general interactive gaming. The intervention was carried out in the acute/subacute stage of stroke during a 12 weeks period. Studies with VR in the acute stage after stroke are rare and little difference in motor function and disability between VR and conventional therapy has been found (Piron et al., 2005). However, taking into account that most of the plastic changes and consequent outcomes happen in the first few months after stroke (Kreisel et al., 2006; Murphy and Corbett, 2009), one would expect that rehabilitation during this period should be more effective. Consequently, it becomes extremely important to investigate whether an early treatment with VR may produce a change in recovery.

Our results suggest that the Rehabilitation Gaming System speeds-up the recovery of the deficits of the upper extremities, with particular emphasis on functional aspects related to the performance of the activities of daily living. This evidences the potential benefits for neurorehabilitation of using VR based systems that directly target the neuronal substrate of recovery through the MNS.

2.. Methods

2.1. Rehabilitation gaming system

The main elements of the Rehabilitation Gaming System (RGS) (Fig. 1) are: the vision based Analysis and Tracking System (AnTS) (Cameirao et al., 2010) that captures upper limb movements through color detection; two data gloves to capture finger flexure (5DT, Fifth Dimension Technologies, Johannesburg, South Africa); an intelligent controller, the Personalized Training Module (PTM) that adapts online the difficulty of the task to the performance of the user; and a virtual environment where an avatar mimics the movements of the user. In the scenario considered here, Spheroids, the user had to interact with approaching flying spheres controlled by parameters such as speed, range of movement and time interval between spheres. These parameters define the difficulty of the task. The training sessions were preceded by two versions of a calibration task, the same task being performed both in the physical and in the virtual environment (Cameirao et al., 2010). In this task patients were asked to move their arms in random sequences to specific positions on the tabletop. In the virtual version, the task was to be performed with the virtual arms moving on a virtual table. This task allowed measuring specific properties of movements such as speed and range of movement, and established the baseline of task difficulty level for every session. Following the baseline calibration, the PTM autonomously defined the baseline difficulty of the Spheroids task. During the training, each new difficulty setting was computed taking into account the previous responses of the user. The difficulty was increased when the user intercepted more than 70% of the spheres; and was decreased if the user intercepted less than 50% of the spheres (Fig. 1). This allowed a continuous adaptation of the game parameters to the user's performance. Moreover, individualization was realized for each arm separately.

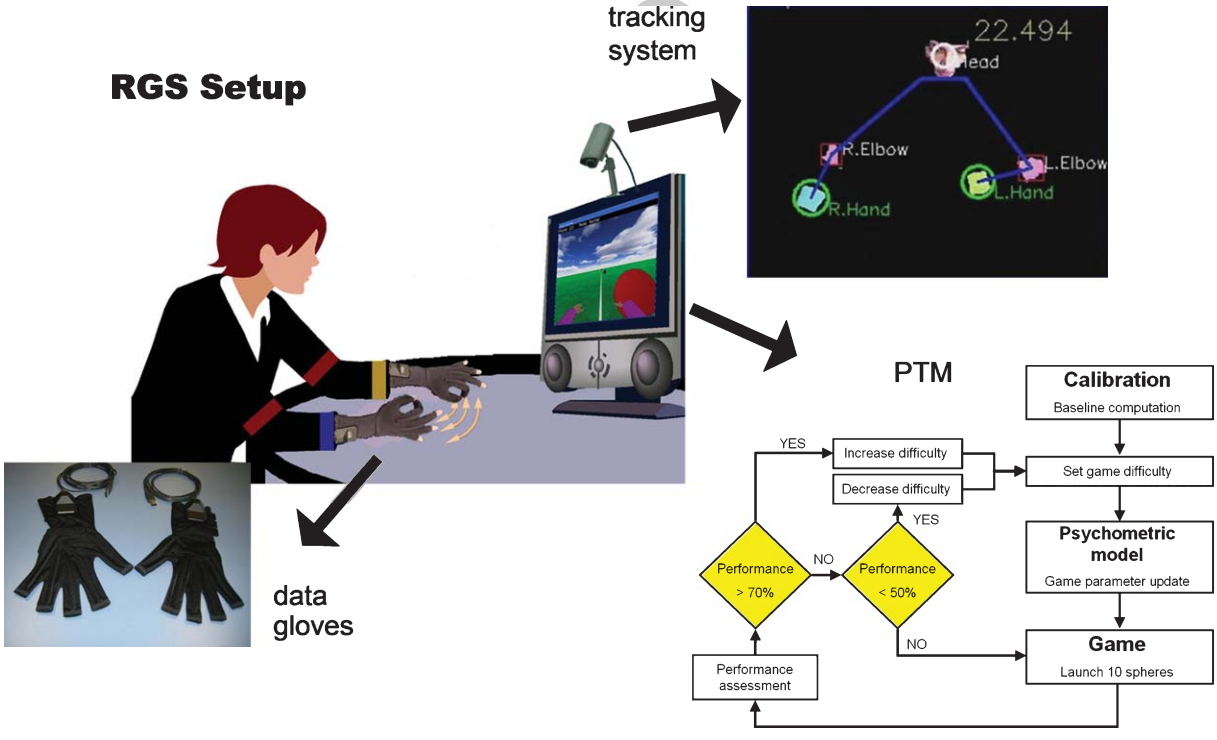

Fig. 1. The Rehabilitation Gaming System. The subject works with his/her arms on a cut-out table facing a computer screen. The movements of the arms are captured by a vision based tracking system that detects color patches positioned on the writs and elbows, and finger bending is measured by data gloves. The captured movements are mapped in real time to the movements of two virtual arms that mimic the movements of the user on the display. In the virtual reality task, Spheroids, approaching spheres have to be intercepted, grasped or placed. The difficulty of the task is adjusted to the user by an intelligent controller, the Personalized Training Module (PTM) that measures the number of successful events and adapts the difficulty accordingly for the next trial (Cameirao, Bermúdez i Badia et al., 2010).

The sessions followed a structured training protocol with tasks of increasing complexity (Hitting, Grasping, and Placing) that train speed and range of movement, grasp and release respectively.

2.2. Subjects and protocol

Subjects were acute stroke patients admitted to the Physical Medicine and Rehabilitation unit of the Hospital de L'Esperança in Barcelona. Out of 142 patients admitted between November 2007 and January 2009, 25 (18%) satisfied the inclusion criteria to participate in the study. The inclusion criteria were: first episode stroke, acute stroke within three weeks post-stroke at baseline, severe to moderate deficit of the paretic upper extremity ($2 \leq \text{MRC} \leq 3$) (MRC, 1976), no severe to moderate aphasia (Rosselli et al., 1990), no other cognitive deficits as assessed by the Mini-Mental State Examination (Folstein et al., 1975), cooperation, and age ≤ 80 years.

After giving their informed consent, patients were randomly assigned to the RGS ($n = 13$) or to the Control group, consisting of either Intense Occupational Therapy (IOT, $n = 6$) or Non-Specific interactive Games (NSG, $n = 6$) using a standard game console. All patients received standard occupational and physical rehabilitation plus the added treatment condition. The patients underwent extended clinical assessment at admittance (baseline), weeks 5, 12 (end of treatment), and 24 (follow-up). The study followed accepted guidelines and was approved by the ethics committee of clinical research of the Parc de Salut Mar.

Out of the original 25 patients selected for the study, one refused to participate and five patients left the study before the week 5 evaluation due to external reasons not related to the treatment (four moved to a different institution and one dropped all rehabilitation). The remaining 19 patients (RGS = 10, 5 males, 63 ± 11 years; Control = 9 (IOT = 5, 3 males, 59 ± 11 years; NSG = 4, 1 male, 58 ± 14 years)) completed the study at least up to week 5 (Table 1). We had missing evaluations for four patients at week 12 and 24: two dropped all rehabilitation half-way the study, one moved to a different institution, and the other one had a second stroke.

2.3. Treatment

In addition to standard rehabilitation, patients had three weekly sessions of 20 minutes each of a given

Table 1
Demographic information of the patients enrolled in the study

Group	ID	Age	Gender	Education	Neurological deficit (NIHSS)	Days after stroke at baseline	Type of stroke	Infarct classification	Side of lesion
RGS	1	79	F	E	13	18	C	TACI	R
	2	60	F	E	4	4	H	-	R
	3	67	M	M	6	9	A	POCI	R
	4	55	M	E	6	13	A	POCI	R
	5	76	M	M	7	16	A	LACI	L
	6	79	F	E	4	7	U	POCI	L
	7	50	F	E	5	8	U	LACI	L
	8	52	M	E	7	19	H	TACI	R
	9	50	F	M	6	13	C	PACI	R
	10	69	M	E	4	8	A	PACI	R
Control IOT	1	66	F	M	7	15	C	LACI	L
	2	54	M	M	8	14	H	-	L
	3	47	M	M	6	22	C	TACI	R
	4	56	M	E	11	11	A	PACI	R
	5	74	F	E	5	22	A	TACI	R
Control NSG	1	65	F	E	2	7	A	LACI	L
	2	37	F	E	6	12	H	-	L
	3	65	M	M	6	18	A	TACI	R
	4	65	F	E	6	15	A	POCI	R

Control: IOT = Intense Occupational Therapy and NSG = Non-Specific Games. Gender: M = male and F = female. Education level: E = elementary and M = medium. Type of stroke (Adams, Bendixen et al., 1993): A = atherosclerotic, C = cardioembolic, H = hemorrhagic and U = undetermined. Infarct classification (Bamford, Sandercock et al., 1991): TACI = total anterior circulation infarct, PACI = partial anterior circulation infarct, POCI = posterior circulation infarct and LACI = lacunar infarct. Lesion side: L = left and R = right.

treatment condition (RGS or Control). Patients in the intervention group performed the Spheroids tasks (Hitting, Grasping, and Placing) introduced gradually during the treatment period. The Control group was split in two subgroups to control different aspects of the intervention. The IOT subgroup carried out pure extended occupational therapy with emphasis on motor tasks similar to the ones promoted by the RGS, namely object displacement, and object grasp and release, but without the action observation component. To control for placebo effects such as computer use and game specific effects, and also for the effect of observing the virtual arms during the task, patients allocated to the NSG subgroup performed games with the Wii system (Nintendo, Tokyo, Japan) that required movements with the paretic arm that did not show any virtual body in response to their actions. i.e., this control had in common with the RGS group the gaming features, but did not share the neuroscientific hypotheses on recovery based on an action observation paradigm. All patients in the Control group performed the RGS calibration task once per week for between-group comparisons.

2.4. Outcome measures

The clinical assessment was performed at baseline, weeks 5, 12, and 24 (follow-up). The evaluators were blind to the assignment of each subject to either the RGS or the Control group. A number of standard clinical scales were used to assess different aspects of motor deficits and function: Barthel Index (Granger, Albrecht et al., 1979) for independence in activities of daily living, Medical Research Council Grade (MRC, 1976) and Motricity Index (Demeurisse et al., 1980) (upper extremities) for muscle strength, Fugl-Meyer Assessment Test (upper extremities) for motor and joint functioning (Fugl-Meyer et al., 1975), and Chedoke Arm and Hand Activity Inventory (CAHAI) (Barreca et al., 2004) for the functional assessment of the paretic arm and hand.

The Rehabilitation Gaming System calibration task allowed us to extract information in terms of speed for both RGS and Control group. In addition, specifically for the RGS group, from the training session, we measured game related events such as successful/unsuccessful trials and difficulty level reached for both the paretic and nonparetic arm.

To assess patients' subjective opinions with respect to a number of aspects of the treatment with RGS such as enjoyment, understanding and ease of the task, patients in the RGS group were given a short self-report questionnaire at the end of the treatment. This questionnaire was presented in the format of a 5-point Likert scale and patients had to report their agreement/disagreement with respect to a number of statements.

2.5. Data analysis

It has been reported that recovery following stroke shows a non-linear logarithmic pattern, with a faster improvement in the first weeks post-stroke followed by smaller improvements at later stages (Kwakkel et al., 2006). In order to correct for this effect, we fitted a logarithmic curve to the individual clinical measures at the different measurement points and assessed the strength of this relation by extracting the squared correlation coefficient, R^2 (Fig. 2). In addition, this logarithmic fit allowed us to estimate missing data, meaning that we had the same number of samples for analysis at each point of measurement. For each scale, for the entire group of patients, we computed the median R^2 and checked the presence of statistical outliers. Median R^2 was of 0.8544 for the Barthel Index, 0.9080 for the Motricity Index, 0.8410 for the upper extremities Fugl-Meyer Score (0.8120 for the arm part and 0.8240 for the wrist/hand part), and 0.8920 for the Chedoke Arm and Hand Activity Inventory. We excluded from the analysis patients that were statistical extreme outliers (values that are more than 3 times the interquartile range above the 75th percentile or below the 25th percentile) in two or more clinical scales. This led to the removal of Patients 5 and 9 in the RGS group and of Patient 2 in NSG subgroup. The dissimilar pattern of recovery of these patients is in accordance with observed personal and clinical circumstances that interfered with the normal progress of these patients during the rehabilitation process.

In order to have an unbiased assessment of the relationship between groups (RGS, IOT and NSG) in the clinical scores, we performed a Principal Components Analysis (PCA) that allowed us to investigate the structure of the data over the groups of patients over all the clinical scales at the end of treatment. We computed the improvements with respect to baseline in the clinical scores at the end of the treatment and standardized the data by dividing each data set by its standard deviation. We then performed the PCA, extracted the principal components scores, and calculate the percent of the total variability explained by each principal

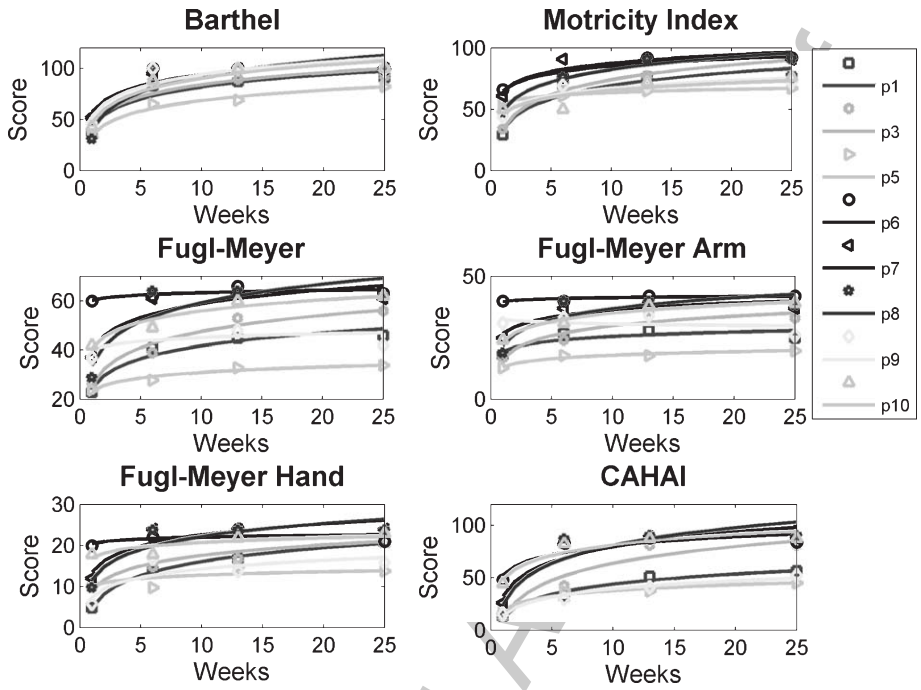

Fig. 2. Time course of improvement at the standard clinical scales. A logarithmic fit has been performed to capture the trend over time. The shown data is from patients in the RGS group with complete clinical evaluations at all time steps.

component. Finally, we performed a between-group comparison on the principal components scores using a two-tailed Mann-Whitney test.

In order to correct for individual differences between participants we computed the *Normalized Improvement* (Eq. 1), which represents the improvement normalized to the total amount that each individual can gain with respect to their baseline.

$$\text{Normalized Improvement}_{i,j} = \left[1 - \frac{(\text{MaxScale}_j - X_i)}{(\text{MaxScale}_j - X_0)}\right] \times 100, \quad i = 1, 2, \ldots \tag{1}$$

where X_i is a given measure of the scale j at time i. X_0 represents baseline.

In order to check balance between groups, the absolute baseline measures were statistically compared using the chi-squared test for categorical data, and a 2-tailed independent samples *t*-test or a Mann-Whitney test for quantitative data. The normality of the distribution was assessed using a single sample Lilliefors hypothesis test of composite normality. To compare the intervention and the control group over time (baseline, end of treatment and follow-up) we performed a repeated measures ANOVA, with time as the within-subject variable and group as the between-subject variable. The between-group comparisons of the normalized improvements at different time points were performed using a 1-tail Mann-Whitney test. For within-group comparisons we used a 2-tail Wilcoxon signed ranks test.

In the analysis of the RGS data, we extracted the weekly average (relative to baseline) of the paretic arm speed in the calibration task for both groups of patients (see methods). The speed time series was smoothed using a moving average with a span of two weeks and to show the trend over time we included a logarithmic fit (see methods). To compare the intervention and the control group over time we performed a Time × Group repeated measures ANOVA, and used a 1-tail Mann-Whitney test for between-group comparisons at time points.

To analyze the evolution of the paretic arm in the Spheroids task, for the RGS group, we extracted the maximum difficulty reached during each session of the Hitting/Grasping task (eight weeks period) and averaged it over periods of two weeks, separately for paretic and nonparetic arms. We computed the difference in difficulty between both arms and removed the statistical outliers at every week (values that are more than

1.5 times the interquartile range above the 75th percentile or below the 25th percentile). We used a 2-tail Wilcoxon signed rank test to compare both arms at each point in time.

Data is expressed as mean ± standard deviation in the text and tables, unless otherwise stated. For all statistical comparisons the significance level was set to 5% ($p < 0.05$). The statistical Power (1-β error probability) as been computed assuming 0.05 alpha and using non-parametric Mann-Whitney tests (Faul, Erdfelder et al., 2007). All statistical analysis was done using MATLAB (version 2008 a) and SPSS (version 16.0).

3. Results

3.1. Outcome measures

In order to have an unbiased assessment of the differences between the RGS group and, the Intense Occupational Therapy and Non-Specific interactive Games control subgroups, we performed a PCA of the clinical improvements at the end of treatment for all groups. The six principal components (PCs) explained 66.21%, 16.03%, 9.30%, 5.00%, 3.44% and 0.01% of the variability of the data, respectively. We observed the existence of a similar recovery pattern for both control interventions, and of a different one for the RGS group. This was particularly salient in the third principal component. The between-group comparisons of the PCs showed no significant differences between the control subgroups for any of the PCs (Mann-Whitney, $p > 0.05$, Power (1-β) ranging from 0.05 to 0.39). However, we found a significant difference between the RGS group and both control subgroups (Mann-Whitney, *RGS-IOT*: $Z = -2.635$, $p < 0{\cdot}01$, $p(1\text{-}\beta) = 0.90$; *RGS-NSG*: $Z = -2.245$, $p < 0.05$, $p(1\text{-}\beta) = 0.64$) for the third PC (Fig. 3). Therefore, taking into account that both control subgroups were statistically indistinguishable from each other while being different from the RGS group, we merged them, the consequent increase of sample size enhancing the statistical power of our analysis. Also further between group comparisons of control subgroups of the improvement at the different clinical scales were not significant (data not shown).

Baseline balance between groups was confirmed for all demographic and clinical measures except for the Fugl-Meyer Assessment Test. The RGS group had a higher score in this measure due to differences in the wrist/hand subpart of the test (Table 2).

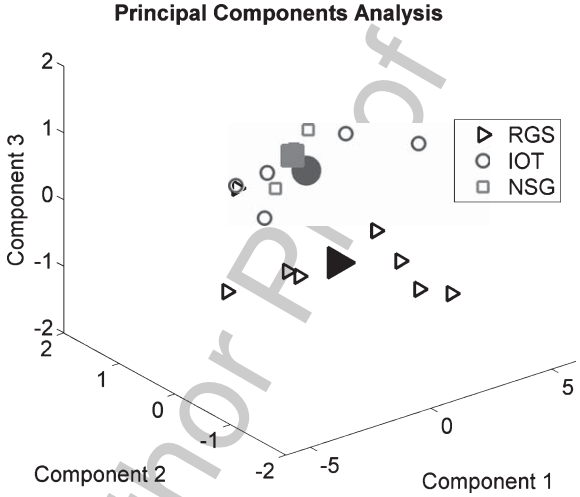

Fig. 3. Structure of the clinical scores of the subjects using a Principal Components Analysis. Representation of the three first Principal Components (PC, 91.56% of variability explained) of the clinical scores at the end of treatment for the RGS group (black), and the control IOT (dark grey) and NSG (light grey) subgroups. The solid markers indicate the centroids (mean) of the distributions for each group.

Table 2
Baseline clinical measures

Variable	RGS ($n = 8$)	Control ($n = 8$)	p-value
Barthel Index (max=100)	42.1 ± 6.8	45.6 ± 14.1	0.537 (T)
MRC (2/3)	4/4	4/4	1.000 (χ^2)
Motricity Index (max=99)	52.2 ± 15.8	42.7 ± 17.7	0.277 (T)
Fugl-Meyer (max=66)	37.9 ± 12.1	24.4 ± 11.4	0.038 (T)
Arm (max=42)	24.8 ± 7.7	18.0 ± 7.1	0.090 (T)
Wrist/Hand (max=24)	13.1 ± 5.0	6.4 ± 4.6	0.015 (M)
CAHAI (max=91)	29.5 ± 15.1	24.5 ± 12.9	0.528 (M)

The categorical variables are expressed in terms of the ratio of cases and the quantitative variables are mean ± standard deviation. For the p-value, the text in brackets denotes the statistical test that was used for the comparison (T = independent samples t-test, M = Mann-Whitney Test, χ^2 = chi squared test).

In the comparison of arm speed between groups in the RGS calibration task, the Time × Group repeated measures ANOVA revealed a significant main effect for Time ($F(3.70, 44.36) = 5.10$, $p < 0.01$, partial eta squared = 0.298) and Group ($F(1, 12) = 6.08$, $p < 0.05$, partial eta squared = 0.336). The Time × Group interaction was leaning towards significance ($F(3.70, 44.36) = 2.59$, $p = 0.053$, partial eta squared = 0.178). Concerning the evolution of speed over time, the RGS showed higher improvements in the paretic arm speed when compared to the control group, and these were

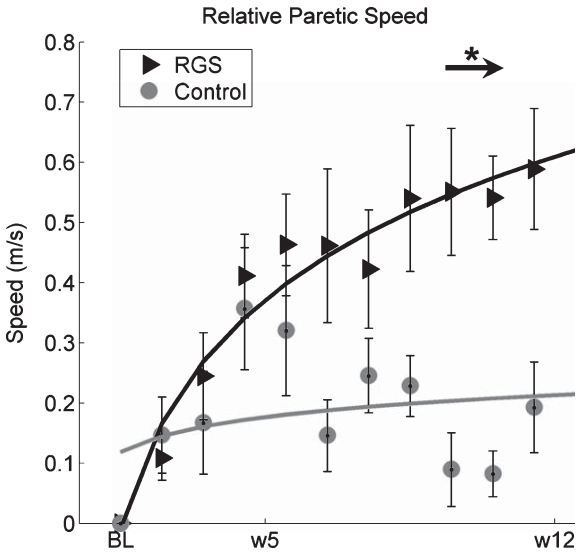

Fig. 4. Speed of the paretic arm over time as measured in the calibration task. Relative average speed (mean ± standard error of the mean) over time for RGS (black) and control (grey) groups. The time series are fitted with logarithmic curves. The arrow indicates the period when the difference between groups starts to be systematically significant, Mann-Whitney test, $*p < 0.05$.

systematically significant after the 9th week of treatment (Mann-Whitney, $p < 0.05$) (Fig. 4). Although the control group showed a steep improvement during the first few weeks, it stabilized after week 5, approximately. This was not the case for the RGS group, which displayed a sustained improvement following a well defined logarithmic pattern ($R^2 = 0.95$).

In the analysis of the specific clinical outcomes assessed by the different clinical scales, the 3(Time) × 2(Group) repeated measures ANOVA revealed a significant main effect for Time for all the clinical measures (*Barthel Index*: $F(1.35, 18.89) = 705.54$, $p < 0.001$, partial eta squared = 0.981; *Motricity Index*: $F(1.45, 20.33) = 205.96$, $p < 0.001$, partial eta squared = 0.936; *Fugl-Meyer*: $F(2, 28) = 177.51$, $p < 0.001$, partial eta squared = 0.927; *Fugl-Meyer Arm subpart*: $F(1.18, 16.52) = 145.31$, $p < 0.001$, partial eta squared = 0.912; *Fugl-Meyer Wrist/Hand subpart*: $F(2, 28) = 96.90$, $p < 0.001$, partial eta squared = 0.874; *CAHAI*: $F(1.32, 18.49) = 388.86$, $p < 0.001$, partial eta squared = 0.965). We found no significant main effect for Group at any measure. However, a significant Time × Group interaction was found for the CAHAI ($F(1.32, 21.13) = 4.09$, $p < 0.05$, partial eta squared = 0.226). In addition, the between-subject comparisons of the normalized improvements at different points in time showed that at the end of the treatment (week 12) the RGS group is significantly better for the arm subpart of the Fugl-Meyer Assessment Test and for the CAHAI, and that this difference was leaning towards significance for the Motricity Index (Table 3). Although the RGS group always showed higher average improvements over time, we found no further significant differences between the groups. Both groups showed significant improvements between baseline and weeks 5 for all the clinical scales. Between weeks 5 and 12, the RGS group improved significantly at all measures (Wilcoxon, *Barthel Index*: $Z = -2.023$, $p < 0.05$, *Motricity Index*: $Z = -2.201$, $p < 0.05$, *Fugl-Meyer*: $Z = -2.201$, $p < 0.05$, *Fugl-Meyer Arm subpart*: $Z = -2.201$, $p < 0.05$, *Fugl-Meyer Wrist/Hand subpart*: $Z = -2.023$, $p < 0.05$, *CAHAI*: $Z = -2.521$, $p < 0.05$), while the control group only improved significantly at the Barthel Index (Wilcoxon, $Z = -2.201$, $p < 0.05$) and the CAHAI (Wilcoxon, $Z = -2.366$, $p < 0.05$). This indicated that the RGS group showed a steeper improvement over time during the treatment period (Fig. 5). No significant improvements were found between week 12 and follow-up for both groups. In summary, the RGS presented on average higher scores at the different points in time, and displayed a sustained faster improvement when compared to the control group.

Finally, we wanted to investigate how accurately the RGS task captured the functional level of the user over time and adjusted the difficulty. The analysis of the maximum difficulty reached over time for the RGS group showed that, as expected, the paretic arm always reached lower levels of difficulty when compared to the nonparetic arm (Fig. 6). However, the paretic arm tended to converge towards the performance of the nonparetic arm during the treatment period. Indeed, the difficulty reached is significantly different between arms at week 2 (Wilcoxon, $Z = -2.380$, $p = 0.05$) and at week 4 (Wilcoxon, $Z = -2.240$, $p < 0.05$), and stopped to be significantly different after the 6 th week of treatment (Wilcoxon, $p > 0.05$). These results show that the RGS captured the functional state of the subject over time and that it autonomously generated the difficulty level accordingly during each session.

3.2. Acceptance and satisfaction

In order to assess the acceptance level of the treatment and the overall satisfaction concerning the use of RGS, patients that performed the entire treat-

Table 3
Normalized improvement at time points

Variable	Week 5			Week 12			Follow-up		
	RGS ($n=8$)	Control ($n=8$)	p/P(1-β)	RGS ($n=8$)	Control ($n=8$)	p/P(1-β)	RGS ($n=8$)	Control ($n=8$)	p/P(1-β)
Barthel	87.6 ± 11.2	81.0 ± 19.4	0.287/0.192	94.9 ± 8.9	88.0 ± 17.8	0.221/0.231	96.3 ± 6.3	92.9 ± 7.1	0.221/0.240
Motricity	52.4 ± 30.0	51.4 ± 22.5	0.253/0.057	73.6 ± 16.1	60.2 ± 20.0	0.052/0.391	81.3 ± 15.9	66.3 ± 20.9	0.065/0.442
Fugl-Meyer	62.0 ± 30.9	55.6 ± 22.1	0.439/0.115	84.6 ± 18.4	66.9 ± 22.9	0.065/0.474	79.1 ± 19.0	72.0 ± 18.8	0.252/0.172
Arm	57.1 ± 36.2	52.9 ± 25.7	0.439/0.081	83.6 ± 19.7	62.3 ± 23.0	0.032/0.597	78.7 ± 24.3	64.6 ± 25.3	0.139/0.277
Wrist/Hand	63.0 ± 36.5	59.1 ± 22.3	0.322/0.080	85.0 ± 21.3	70.6 ± 32.2	0.191/0.253	85.7 ± 25.5	81.5 ± 12.7	0.080/0.104
CAHAI	72.7 ± 26.5	46.5 ± 29.6	0.065/0.533	90.2 ± 17.0	70.6 ± 18.2	0.025/0.662	89.6 ± 14.9	81.9 ± 12.3	0.080/0.275

The normalized improvements are expressed as mean ± standard deviation. A 1-tail Mann-Whitney test was used for the statistical comparisons. P(1-β) is the Power (1- β error probability) assuming 0.05 alpha and 1-tail Mann-Whitney test.

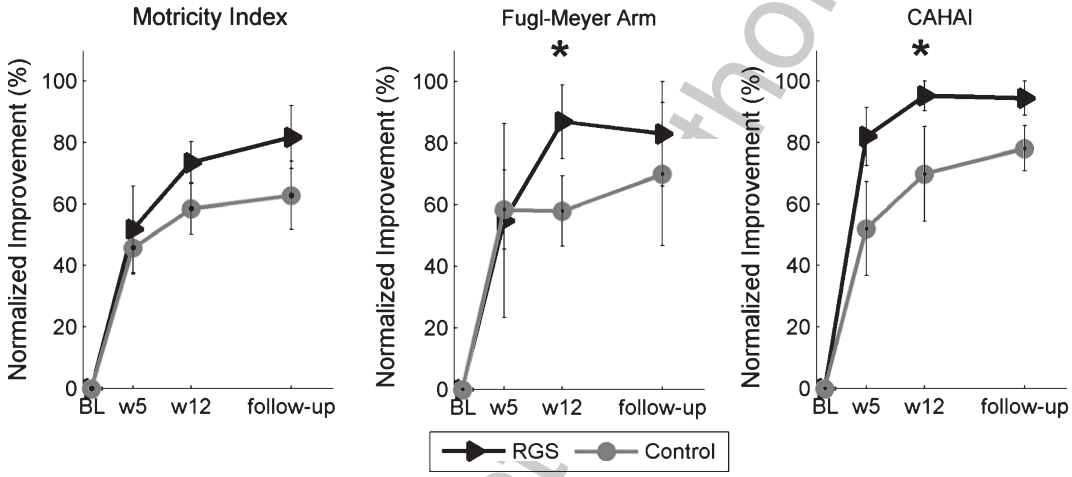

Fig. 5. Normalized improvement over time for the Motricity Index, the arm subpart of the Fugl-Meyer Assessment Test and the Chedoke Arm and Hand Activity Inventory. Improvement (median ± median absolute deviation) for RGS (black) and control (gray) groups for selected clinical scales. *$p<0.05$, between-group comparison.

ment period with RGS ($n=8$) were given a succinct self-report 5-point Likert questionnaire at the end of treatment. Statements could be rated as 1 (strongly disagree), 2 (disagree), 3 (neither agree nor disagree), 4 (agree) or 5 (strongly agree). This allowed us assessing a number of aspects such as enjoyment, understanding and ease of the task. In addition, patients were also asked if they would like to continue the treatment with RGS. In terms of enjoyment, to the statement "The task was entertaining", the average rating was 4.5. To the statement "The task was too long", the average rating was 1.2. In terms of clarity and difficulty in using the system, to the statement "The task was easy to understand", the average rating was 4.9. To the statement "It was difficult to control the virtual arms", the average rating was 2.1. Finally, as a measure of overall satisfaction, to the statement "I would like to continue this treatment", the average rating was 4.4. Based on these results and as an overall analysis we feel confident to conclude that the acceptance of the RGS and its tasks was very high.

4. Discussion

We have investigated the impact of a novel VR based rehabilitation paradigm, the Rehabilitation Gaming Station, on the functional recovery of deficits of the upper extremities of acute/subacute stroke patients. Our results indicate that the RGS group followed a substantially different pattern of recovery when compared to the control group. We observed that at the end of the treatment the RGS group performed better as compared to controls in both the speed of the paretic arm and the scores on a number of clinical scales. In addition, the RGS group presented a significantly

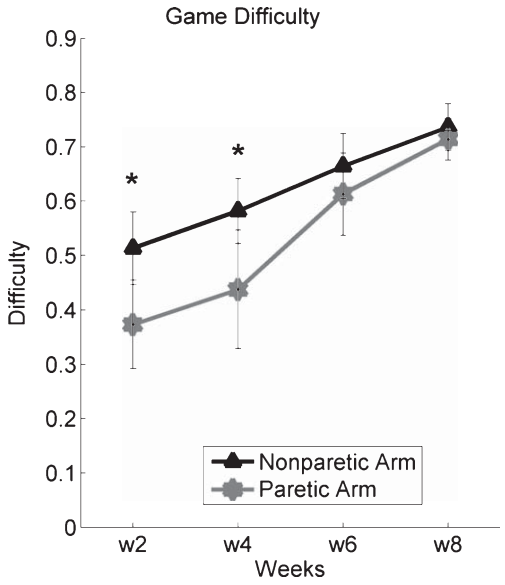

Fig. 6. Game difficulty. Biweekly average of the difficulty level reached during the Hitting/Grasping task in the Spheroids game (mean ± standard error of the mean) for paretic (grey) and non-paretic (black) arms. The difficulty level goes up to a maximum of 1.0. $*p < 0.05$, pairwise comparison.

faster improvement over time for all the clinical scales during the treatment period. Specifically, in the evolution over weeks of the average paretic arm speed in the RGS calibration task, the RGS group showed in general a higher movement speed when compared to the control group, and there was a statistically significant difference after the 9th week of treatment. This could be related to the fact that higher arm speed was required in order to accomplish higher difficulty levels in the RGS tasks. Therefore RGS patients were indirectly developing higher movement speed skills. In the analysis of detailed clinical outcomes assessed by standard clinical evaluation at the different time stages, patients allocated to the RGS group showed in general higher improvements and these were particularly salient at the end of the treatment. Specifically, the between group difference was statistically significant for the arm subpart of the Fugl-Meyer Assessment Test and for the Chedoke Arm and Hand Activity Inventory. Hence, RGS supported rehabilitation seems to have a particular impact on the recovery of proximal movements and on the ability to perform functional activities of daily living. Since the RGS promotes proximal and distal movements, we would also expect to have a significant impact at the hand subpart of the Fugl-Meyer Test, but this was not the case. We have two possible explanations for this. First, it could be due to imbalance at baseline for this specific measure. Second, although RGS trains finger grasping and release, there are only virtual objects to be grasped and the patient had no physical contact with them. Therefore, there was no sensory information on the effectiveness of this movement. This may indicate the need to incorporate a graspable object preferably coupled with a haptic interface to provide sensorimotor feedback and increase the ecological validity of the task (Levin et al., 2009). To address this issue, we are currently developing an updated version of the RGS that integrates a haptic interface that provides sensorimotor feedback during the task.

Newer technology driven rehabilitation strategies such as robotics, functional electrical stimulation and transcranial magnetic stimulation have shown so far good outcomes at the movement level but with poor outcomes at the functional performance of activities of daily living (ADL) (Mehrholz et al., 2008; O'Dell et al., 2009). In contrast, in our study the RGS group showed a considerable improvement at the performance of ADLs, as measured by the Chedoke arm and hand activity inventory. We believe that the main contributing factor of RGS to this functional impact is its theoretical rationale that aims at tackling the central nervous system, as opposed to emphasizing the manipulation of the peripheral skeleton-motor system. However, given the lack of additional imaging data to confirm this fact, this has to be interpreted with caution as we cannot exclude other potentially beneficial factors such as the effect of treatment personalization and the adaptive nature of the system to sustain performance and motivation. In addition, further experiments are required to fully assess this.

The clinical scores over time showed that, although we observed significant group differences at the end of treatment, this significance was lost at follow-up (12 weeks after the end of the treatment). This could mean that rehabilitation with RGS predominantly accelerates recovery following stroke. Indeed, our results showed that only the RGS group improved significantly at all clinical scales, systematically from baseline to week 5 of treatment and from week 5 to end of treatment. i.e., the RGS group presented a steeper improvement over time during the treatment period. On the basis of this result it is important to investigate if the RGS just speeds-up recovery or if it could more markedly enhance recovery if we increase the intensity of the treatment and/or the longitudinal time duration of the

intervention. We are currently running clinical trials that address the relationship between treatment intensity and duration. In addition, it is important to further assess the impact of VR on the early stages of stroke. Most plastic changes occur during this period and therefore recovery could be possibly maximized (Murphy and Corbett, 2009).

Finally, we showed that the RGS was able to capture the functional dissimilarities between paretic and nonparetic arms and adapted the difficulty of the task accordingly. In this way we provide an autonomous adaptable training regime that is directed towards the individual needs and capabilities of the patients. In addition, this results in higher levels of motivation and compliance with the treatment as shown by the results of our acceptance study. Indeed, the opinion of the patients that used the RGS shows that the majority would like to continue therapy with the RGS.

Our results indicate that rehabilitation with the Rehabilitation Gaming System facilitates the functional recovery of the upper extremities in the acute phase of stroke. Although our results are exciting, this study has as limitation the small size of the sample and further testing is needed with larger populations of patients. Moreover, brain imaging methods should be used to assess the specific benefits of RGS at the level of cortical reorganization. Despite these limitations, our results show promise in terms of the benefits provided by the RGS for the neurorehabilitation of motor deficits following stroke.

Acknowledgments

This work was supported by the European Commission through the projects PRESENCCIA (IST-2006-27731) and RGS (EC-FP7 Ambient Assisted Living Joint Programme, AAL-2008-1-119).

References

Adams, H.P. Jr, Bendixen, B.H. et al. (1993). Classification of subtype of acute ischemic stroke. Definitions for Use in a Multicenter Clinical Trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke*, *24*(1), 35-41.

Bamford, J., Sandercock, P. et al. (1991). Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet 337*(8756), 1521-1526.

Bao, S., Chan, V.T. et al. (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature 412*(6842), 79-83.

Barreca, S., Gowland, C.K. et al. (2004). Development of the chedoke arm and hand activity inventory: theoretical constructs, item generation, and selection. *Top Stroke Rehabil*, *11*(4), 31-42.

Cameirao, M.S., Bermúdez i Badia, S. et al. (2010). Neurorehabilitation using the virtual reality based rehabilitation gaming system: methodology, design, psychometrics, usability and validation. *J Neuroeng Rehabil* 7, 48.

Cameirao, M.S., Bermúdez i Badia, S. et al. (2008). Virtual reality based upper extremity rehabilitation following stroke: a review. *Journal of CyberTherapy & Rehabilitation 1*(1), 63-74.

Demeurisse, G., Demol, O. et al. (1980). Motor evaluation in vascular hemiplegia. *Eur Neurol 19*(6), 382-389.

Faul, F., Erdfelder, E. et al. (2007). G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods 39*(2), 175-191.

Folstein, M.F., Folstein, S.E. et al. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res 12*(3), 189-198.

Fugl-Meyer, A.R., Jaasko, L. et al. (1975). The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. *Scand J Rehabil Med 7*(1), 13-31.

Granger, C.V., Albrecht, G.L. et al. (1979). Outcome of comprehensive medical rehabilitation: measurement by PULSES profile and the barthel index. *Arch Phys Med Rehabil 60*(4), 145-154.

Green, C.S. & Bavelier, D. (2008). Exercising your brain: a review of human brain plasticity and training-induced learning. *Psychol Aging 23*(4), 692-701.

Kalra, L. (2009) Stroke rehabilitation: old chestnuts and new insights. *Stroke 41*(2): e88-e90.

Krakauer, J.W. (2006). Motor learning: its relevance to stroke recovery and neurorehabilitation. *Curr Opin Neurol 19*(1), 84-90.

Kreisel, S.H., Bazner, H. et al. (2006). Pathophysiology of stroke rehabilitation: temporal aspects of neuro-functional recovery. *Cerebrovasc Dis 21*(1-2): 6-17.

Kwakkel, G., Kollen, B. et al. (2006). Impact of time on improvement of outcome after stroke. *Stroke 37*(9), 2348-2353.

Levin, M.F., Knaut, L.A. et al. Virtual reality environments to enhance upper limb functional recovery in patients with hemiparesis. *Stud Health Technol Infor 145*, 94-108.

Lucca, L.F. (2009). Virtual reality and motor rehabilitation of the upper limb after stroke: a generation of progress? *J Rehabil Med 41*(12): 1003-100.

Mathers, C.D. & Loncar D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med 3*(11): e442.

Mehrholz, J., Platz, T. et al. (2008). Electromechanical and robot-assisted arm training for improving arm function and activities of daily living after stroke. *Cochrane Database Syst Rev* (4): CD006876.

MRC (1976). Medical Research Council of the UK. Aids to the Investigation of Peripheral Nerve Injuries. London, Pendragon House.

Mukamel, R., Ekstrom, A.D. et al. (2010). Single-neuron responses in humans during execution and observation of actions. *Curr Biol*.

Murphy, T.H. & Corbett D. (2009). Plasticity during stroke recovery: from synapse to behaviour. *Nat Rev Neurosci 10*(12), 861-872.

O'Dell, M.W., Lin, C.C. et al. (2009). Stroke rehabilitation: strategies to enhance motor recovery. *Annu Rev Me 60*, 55-68.

O'Regan, J.K. & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav Brain Sci 24*(5): 939-73; Discussion 973-1031.

Olshausen, B.A. & Field, D.J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature 381*(6583): 607-609.

Piron, L., Tonin, P. et al. (2005). Virtual environment training therapy for arm motor rehabilitation. *Presence 14*(6), 732-740.

Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nat Rev Neurosci 6*(7), 576-582.

Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annu Rev Neurosc 27*, 169-192.

Rizzolatti, G. & Fabbri-Destro, M. (2010). Mirror neurons: from discovery to autism. *Exp Brain Res 200*(3-4): 223-237.

Rosselli, M., Ardila, A. et al. (1990). Normative data on the boston diagnostic aphasia examination in a spanish-speaking population. *J Clin Exp Neuropsychol 12*(2), 313-322.

Sanchez-Montanes, M.A., Verschure, P.F. et al. (2000). Local and global gating of synaptic plasticity. *Neural Comput 12*(3), 519-529.

Ungerleider, L.G., Doyon, J. et al. (2002). Imaging brain plasticity during motor skill learning. *Neurobiol Learn Mem 78*(3), 553-564.

Verschure, P.F., Voegtlin, T. et al. Environmentally mediated synergy between perception and behaviour in mobile robots. *Natur 425*, 620-624.

Wyss, R., Konig, P. et al. (2006). A model of the ventral visual system based on temporal stability and local memory. *PLoS Biol 4*(5), e120.