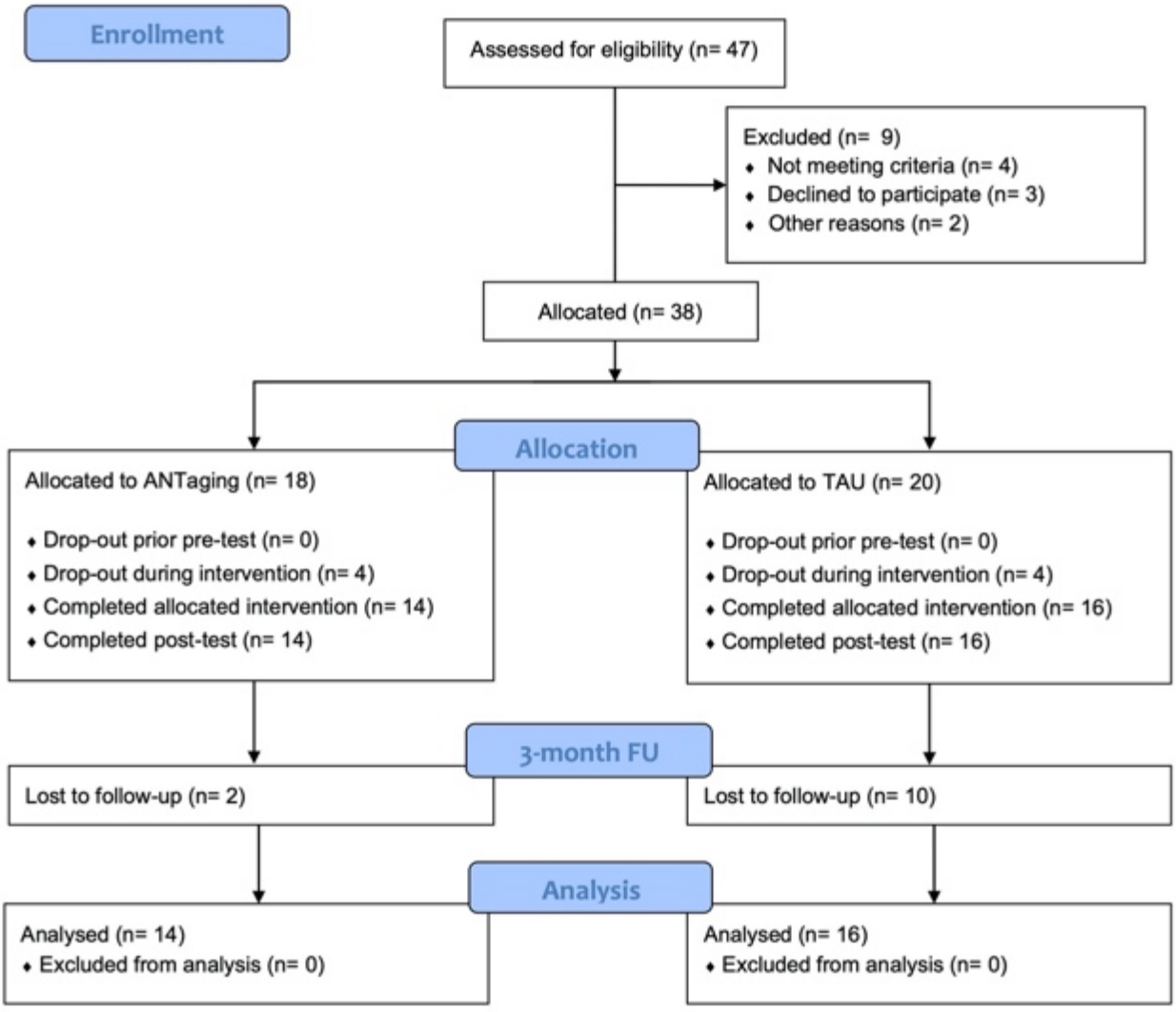
Supplementary Material

Embodied Spatial Navigation Training in Mild Cognitive Impairment: A Proof-of-Concept Trial

Figure 1. CONSORT flow-chart



Neuropsychological battery details

The Mini-Mental State Examination (MMSE) is a brief global cognition test that is used to exclude cognitive deterioration (scores 0-30; greater score better performance). We used the Italian MMSE validation by Magni and colleagues [1] with a correction by age and education level. The forward Corsi block tapping test (CBT) is a test used to assess visuospatial short-term memory (scores 3-9; greater score better performance). We employed the Italian CBT validation by Monaco and colleagues [2] with a correction by age and education level. The digit span forward (DSF) is a test used to assess auditory-verbal short-term memory (scores 3-9; greater score better performance). We employed the Italian CBT validation by Monaco and colleagues [2] with a correction by age and education level. The story recall (SR) is a test used to assess immediate (SR-I) and delayed recall (SR-D) (scores 0-8; greater score better performance). In addition, it is possible to compute SR forgetting (SR-F) by subtracting SR-I minus SR-R (scores 0-8; greater score higher forgetting). We employed the Italian SR validation by Carlesimo and colleagues [3] with a correction by age and education level. To assess visuospatial long-term memory, we used the Corsi supra-span (CSS) (score 0.06-29.16; greater score better performance). We employed the Italian CSS validation by Spinnler and Tognoni [4] with a correction by age, education level, and CBT span. For spatial mental rotation, we used the Manikin test (MT) (scores 0-1; greater score better performance). We employed the standard version [5]. In this test, the patient is presented with an upward or downward and front or back-facing manikin holding a black circle, the participant must tell in which hand the manikin is holding the black circle. In addition to the MT total score, we computed derived scores depending on the manikin orientation: manikin front upward (Mfu), manikin back upward (Mbu), manikin front downward (Mfd), manikin back downward (Mbd). MT is considered a task of egocentric transformation, where participants had to mentally change their egocentric coordinates to a new set of egocentric ones to answer correctly [6]. This change is sustained by additional allocentric processing that supports egocentric translocation [7]. Therefore, the Mbu condition requires no egocentric transformation, the Mbd requires no egocentric transformation but a vertical egocentric rotation (downward to upward), the Mfu requires egocentric transformation, and the Mfd requires egocentric transformation plus a vertical egocentric rotation.

Executive functions were assessed using the Trail Making Test (part A, part B, part BA; TMT-AB). Psychomotor speed was assessed with TMT-A (0-300; greater score worse

performance), divided attention with TMT-B (0-300; greater score worse performance), and central executive with TMT-BA (0-300; greater score worse performance). We employed the Italian version of the TMT-AB [8] with correction by age and education.

ANTaging Encoding and Recall instructions and details

Before the start of the encoding phase, the experiment read an instruction concerning the task to be carried out.

Instructions for the immediate recall phase varied according to the session number. From session one to five the virtual metric was influenced by the fading cues at retrieval (i.e., correct location marker); from session six to ten, responses represented pure patient ability based solely on the feedback provided by the software. During recall, any cues (directional lines, attentional markers or maps) were provided to the patients and during the whole training, a verbal interference (read a two-page story) task was administered between the encoding and recall phases.

Encoding instructions

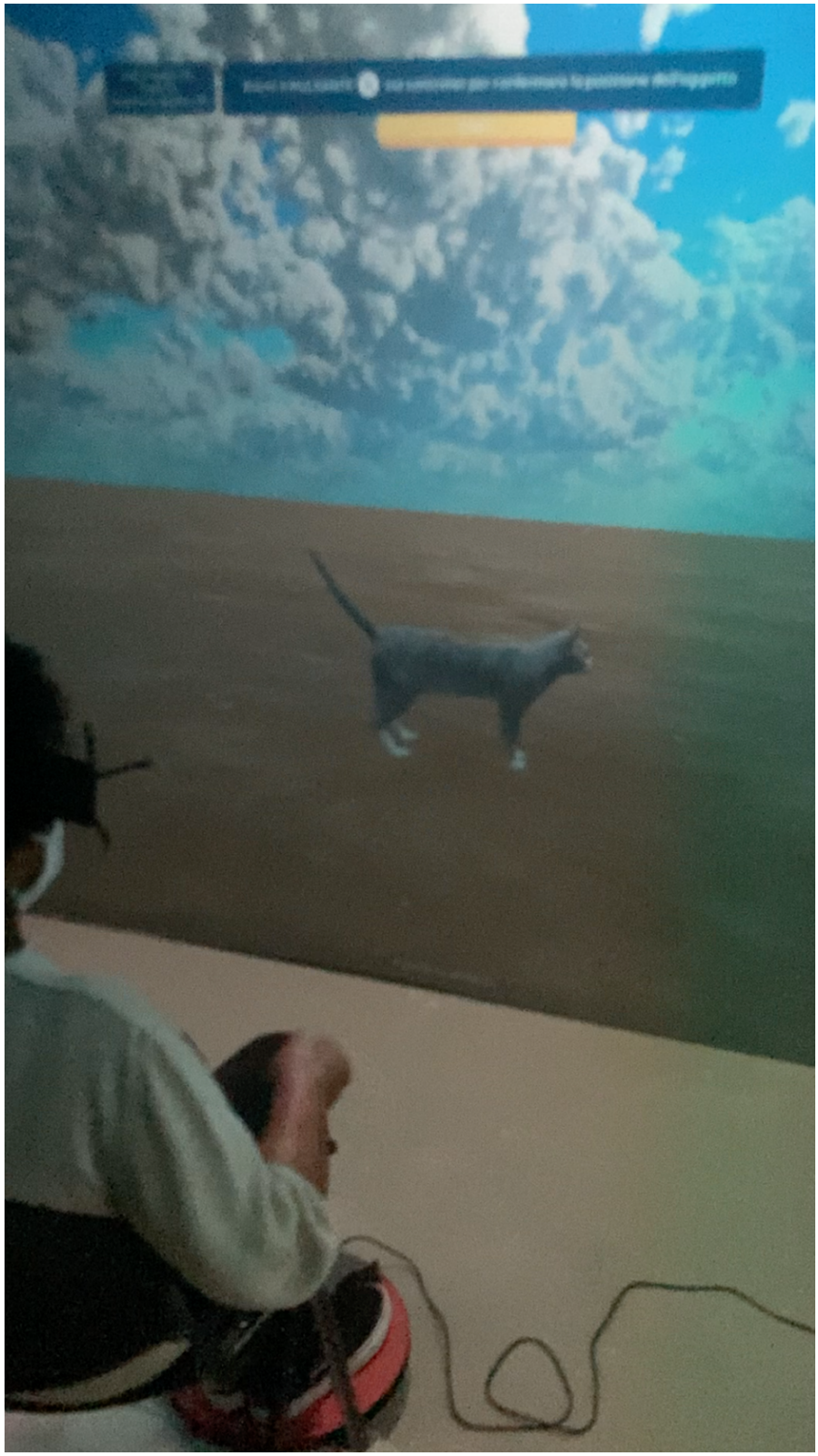
"*Now you will be in a circular city square, and your task is to collect some objects and memorize their locations because you will then be asked to remember them later. You will see one object at a time. You will see each object four times in the same position so that you can remember better its location. To help you to memorize the location you can use the obelisk, the arcade of the square, the mountain range, and the fixed clouds as references. To collect the object, go exactly over it. It will disappear and you will be presented with the next one*".

Recall instructions

For the first two sessions, instructions of the recall phase were "*You will be asked to put each object in the position where you collected it. Your task is to look for a white circle* (correct position marker) *in the city square, which represents the exact location of the object you need to relocate. Once you are above the white circle, press the button A on the joypad. You will be asked to replace each object several times regardless of the correctness of your answer. Note that in random order either the obelisk or the wall will be removed.*".

For sessions number three to five, recall instructions were modified. "*You will be asked to put each object in the position where you collected it. Your task is to look for the position you think is correct, once happy with the location or if you do not remember it, you will be shown the white circle which represents the correct location of the item you have to relocate. Once you are above the white circle, press the button A on the joypad. You will be asked to replace each object several times regardless of the correctness of your answer. Note that in random order either the obelisk or the wall will be removed*".

For the last five sessions, recall instructions were modified again. "*You will be asked to put each object in the position where you collected it. Your task is to look for the position you think is correct, once happy press the button A on the joypad. Written feedback will tell you if you placed the item in the correct location or to try again. You will be asked to replace each object several times regardless of the correctness of your answer. Note that in random order either the obelisk or the wall will be removed*". The correct written feedback ("*Well done! This is the correct location*") appeared if the response was within a radius of 6 virtual units from the actual location, otherwise a try again essage was displayed ("*Try again, you are too far*").



Note 1. A patient during the ANTaging recall sessions (no boundary; egocentric recall). See the 3dRudder under the feet. Written consent for this photo was obtained from the patient.

Pre-post ANOVA

Due to the imbalanced missing rate in the follow-up between groups, we also conducted a mixed (2x2) ANOVA adjusted for TMT-A and SR-F baseline scores for the MT and CSS models. Time had two within-subjects levels (pre-test versus post-test) and group had two between-subjects levels (TAU versus ANTaging). For the CSS model, we found again a significant main effect of Time ($F_1 = 8.12$, $p = 0.005$, $\eta^2_p = 0.13$). Regardless of the intervention CSS improved from pre-test to post-test (est. diff. = 0.32, SE = 0.97). No other effect or contrast was significant. CSS was replicated from the 3x2 model. Conversely, we did not find any significant results for the MT model. The 3x2 model MT was not replicated.

We also conducted a mixed (2x2) ANOVA adjusted for TMT-A and SR-F baseline scores for the relevant secondary outcomes. Time had two within-subjects levels (pre-test versus post-test) and group had two between-subjects levels (TAU versus ANTaging). Only three models had significant results on the independent variables.

Concerning the MMSE model, we found a Time by Group significant interaction ($F_1 = 5.68$, $p = 0.024$, $\eta^2_p = 0.17$). No other effect was significant. Nevertheless, post-hoc contrasts were not significant. The MMSE result found in the 3x2 model was not replicated.

Concerning the SR-F, we found a significant Time by Group interaction ($F_1 = 6.01$, $p = 0.018$, $\eta^2_p = 0.10$). Post-hoc comparisons showed that in the ANTaging group scores improved from pre-test to post-test (est. diff. = -0.73, SE = 0.26, $p = 0.006$). Any other contrasts were found to be significant. The SF-R results in the 3x2 model were replicated.

Regarding the Mbu, we found a significant Time by Group interaction ($F_1 = 5.67$, $p = 0.023$, $\eta^2_p = 0.17$). No other effect was found to be significant. Post-hoc contrast showed that in the TAU group scores declined from pre-test to post-test (est. diff. = -0.14, SE = 0.06, $p = 0.044$). Any other contrasts were found to be significant. The Mbu finding in the 3x2 model was not replicated.

Table 1. Longitudinal clinical significance (Cohen's d) within groups

	TAU, N = 16				ANTaging, N = 14			
Outcome	**Pre**	**Post**	**Paired Cohen's d**	**95% CI**	**Pre**	**Post**	**Paired Cohen's d**	**95% CI**
MMSE (points)	25.80 (2.88)	26.93 (2.49)	-0.48	-0.99, 0.05	25.95 (3.01)	24.71 (3.41)	0.4	-0.15, 0.94
SR-I (points)	4.11 (2.65)	5.21 (2.21)	-0.34	-0.84, 0.17	4.61 (2.53)	4.66 (2.32)	-0.02	-0.55, 0.50
SR-D (points)	4.68 (3.24)	5.34 (2.16)	-0.17	-0.66, 0.33	3.50 (2.99)	4.06 (2.95)	-0.31	-0.84, 0.23
SR-F (points)	0.04 (0.10)	0.13 (0.27)	-0.15	-0.69, 0.40	1.23 (1.60)	0.50 (1.40)	0.36	-0.26, 0.97
CBT (points)	4.05 (1.23)	4.29 (0.75)	-0.19	-0.68, 0.31	4.24 (0.78)	4.53 (0.83)	-0.31	-0.84, 0.23
CSS (points)	7.5 (3.2)	9.4 (5.4)	-0.38	-0.89, 0.13	8.4 (4.5)	10.9 (5.7)	**-0.81**	**-1.41, -0.19**
DSF (points)	5.37 (0.98)	5.55 (1.27)	-0.17	-0.66, 0.33	5.61 (0.81)	5.90 (0.91)	-0.48	-1.02, 0.09
TMT-A (s)	71 (34)	61 (41)	0.32	-0.19, 0.82	43 (14)	55 (47)	-0.32	-0.86, 0.22
TMT-B (s)	172 (103)	157 (106)	0.17	-0.33, 0.66	154 (141)	153 (151)	0.02	-0.51, 0.54
TMT-BA (s)	109 (83)	97 (90)	0.14	-0.35, 0.63	115 (130)	99 (110)	0.23	-0.30, 0.76
MT (accuracy)	0.70 (0.23)	0.68 (0.25)	0.11	-0.38, 0.60	0.69 (0.24)	0.75 (0.19)	-0.55	-1.10, 0.02
Mfu (accuracy)	0.65 (0.34)	0.73 (0.32)	-0.30	-0.80, 0.21	0.64 (0.39)	0.67 (0.31)	-0.08	-0.61, 0.44
Mbu (accuracy)	0.79 (0.22)	0.65 (0.33)	**0.61**	**0.07, 1.14**	0.74 (0.29)	0.83 (0.18)	-0.31	-0.84, 0.23
Mfd (accuracy)	0.5 (0.38)	0.56 (0.36)	-0.19	-0.68, 0.31	0.68 (0.39)	0.61 (0.36)	0.17	-0.36, 0.70
Mbd (accuracy)	0.80 (0.33)	0.80 (0.32)	0.00	-0.49, 0.49	0.71 (0.35)	0.84 (0.29)	-0.43	-0.97, 0.13

MMSE, Mini-Mental State Examination; SR-I, story recall immediate; SR-D, story recall delayed; SR-F, story recall forgetting; CBT, Corsi block tapping test; CSS, Corsi supra-span test; DSF, digit span forward; TMT-AB, Trail Making Test (part A, part B, part BA) MT, manikin test; Mfu, manikin front upward; Mbu, manikin back upward; Mfd, manikin front downward; Mbd, manikin back downward. Bold values show significant results.

Table 3. Between groups clinical significance (Cohen's d) at post-test

Outcome	TAU, N = 16	VR, N = 14	Cohen's d	95% CI
MMSE (points)	26.93 (2.49)	24.71 (3.41)	0.75	0.00, 1.5
SR-I (points)	5.21 (2.21)	4.66 (2.32)	0.24	-0.48, 0.96
SR-D (points)	5.34 (2.16)	4.06 (2.95)	0.50	-0.24, 1.2
SR-F (points)	0.13 (0.27)	0.50 (1.40)	-0.40	-1.2, 0.39
CBT (points)	4.29 (0.75)	4.53 (0.83)	-0.30	-1.0, 0.42
CSS (points)	9.4 (5.4)	10.9 (5.7)	-0.26	-0.97, 0.47
DSF (points)	5.55 (1.27)	5.90 (0.91)	-0.32	-1.0, 0.41
TMT-A (s)	61 (41)	55 (47)	0.12	-0.59, 0.84
TMT-B (s)	157 (106)	153 (151)	0.03	-0.68, 0.75
TMT-BA (s)	97 (90)	99 (110)	-0.01	-0.73, 0.70
MT (accuracy)	0.68 (0.25)	0.75 (0.19)	-0.28	-1.0, 0.44
Mfu (accuracy)	0.73 (0.32)	0.67 (0.31)	0.18	-0.54, 0.90
Mbu (accuracy)	0.65 (0.33)	0.83 (0.18)	-0.65	-0.84, 0.60
Mfd (accuracy)	0.56 (0.36)	0.61 (0.36)	-0.12	-0.84, 0.60
Mbd (accuracy)	0.80 (0.32)	0.84 (0.29)	-0.14	-0.86, 0.58

MMSE, Mini-Mental State Examination; SR-I, story recall immediate; SR-D, story recall delayed; SR-F, story recall forgetting; CBT, Corsi block tapping test; CSS, Corsi supra-span test; DSF, digit span forward; TMT-AB, Trail Making Test (part A, part B, part BA); MT, manikin test; Mfu, manikin front upward; Mbu, manikin back upward; Mfd, manikin front downward; Mbd, manikin back downward. Bold values show relevant results.

aMCI analyses (N = 28; ANTaging = 13, TAU = 15)

We reported the significant results of the 15 ANOVA models (2x3 ANOVAs adjusted for baseline TMT-A and SF-R) for the aMCI sample.

Concerning MT, we only found a statistical tendency for the group by time interaction (F_2 = 2.72, p = 0.07, η^2_p = 0.09).

Concerning CSS, we found a main effect of time (F_2 = 4, p = 0.024, η^2_p = 0.13). SF-R baseline was significant (F_1 = 5.44, p = 0.028, η^2_p = 0.18). CSS improved at FU compared to the pre-test (est. diff. = 2.79, SE = 1.02, p = 0.026).

We found a group by time interaction for the MMSE (F_2 = 4.03, p = 0.024, η^2_p = 0.13). No other effect was found. In the TAU group we found better scores for post-test compared to FU (est. diff. = 1.75, SE = 0.69, p = 0.043). No results were found for the ANTaging group.

We found a group by time interaction for the SF-R (F_2 = 6.43, p = 0.003, η^2_p = 0.2). Baseline SF-R was significant (p < 0.001). In the ANTaging group, forgetting improved after the intervention compared to pre-test (est. diff. = -0.795, SE = 0.28, p = 0.018). In the TAU group

forgetting worsened from pre-test to FU (est. diff. = 0.678, SE= 0.26, p = 0.034). No other effect was found.

We found a significant group by time interaction for the Mbu score ($F_2 = 3.53$, $p = 0.036$, $\eta^2_p = 0.12$). However, post-hoc did not show significant contrasts. No other effect was found.

Regarding pre versus post-test clinical significance in the two interventions, results were unchanged from the whole MCI sample. Regarding differences between the two interventions at post-test, we found a large beneficial change at post-test for the Mbu score in the ANTaging group (d = -0.80, 95%CI [-1.56, -0.02]) compared to TAU. This was not found in the MCI sample.

Regarding egocentric and allocentric performance across the last five sessions in the aMCI sample allocated to ANTaging, results were the same as the whole MCI group.

Overall, results mainly replicate MCI whole sample findings.

REFERENCES

[1] Magni E, Binetti G, Bianchetti A, Rozzini R, Trabucchi M (1996) Mini-mental state examination: A normative study in Italian elderly population. *Eur J Neurol* **3**, 198–202.

[2] Monaco M, Costa A, Caltagirone C, Carlesimo GA (2012) Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an Italian adult population. *Neurol Sci* **34**, 749–754.

[3] Carlesimo GA, Buccione I, Fadda L, Graceffa A, Mauri M, Lorusso S, Bevilacqua G, Caltagirone C (2002) Stan-dardizzazione di due test di memoria per uso clinico. Breve racconto e figura di Rey. *Nuova Riv di Neurol* **12**, 1–13.

[4] Spinnler H, Tognoni G (1987) Standardizzazione e taratura italiana di test neuropsicologici. *Ital J Neurol Sci* **8**, 1–120.

[5] Ratcliff G (1979) Spatial thought, mental rotation and the right cerebral hemisphere. *Neuropsychologia* **17**, 49–54.

[6] Kaltner S, Riecke BE, Jansen P (2014) Embodied mental rotation: a special link between egocentric transformation and the bodily self. *Front Psychol* **5**, 505.

[7] Vogeley K, Fink GR (2003) Neural correlates of the first-person-perspective. *Trends Cogn Sci* **7**, 38–42.

[8] Mondini S, Mapelli D, Vestri A, Bisiacchi PS (2003) *Esame neuropsicologico breve. Una batteria di test per lo screening neuropsicologico.*