

ORIGINAL ARTICLE

Validation of the Mnemonic Similarity Task – Context Version

Giulia A. Aldi,¹ Iris Lange,² Cristiana Gigli,¹ Lies Goossens,² Koen R. Schruers,² Fiammetta Cosci¹

¹Dipartimento di Scienze della Salute, Università di Firenze, Firenze, Italy. ²Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, The Netherlands.

Objective: Pattern separation (PS) is the ability to represent similar experiences as separate, non-overlapping representations. It is usually assessed via the Mnemonic Similarity Task – Object Version (MST-O) which, however, assesses PS performance without taking behavioral context discrimination into account, since it is based on pictures of everyday simple objects on a white background. We here present a validation study for a new task, the Mnemonic Similarity Task – Context Version (MST-C), which is designed to measure PS while taking behavioral context discrimination into account by using real-life context photographs.

Methods: Fifty healthy subjects underwent the two MST tasks to assess convergent evidence. Instruments assessing memory and attention were also administered to study discriminant evidence. The test-retest reliability of MST-C was analyzed.

Results: Weak evidence supports convergent validity between the MST-C task and the MST-O as measures of PS ($r_s = 0.464$; $p < 0.01$); PS performance assessed via the MST-C did not correlate with memory or attention; a moderate test-retest reliability was found ($r_s = 0.595$; $p < 0.01$).

Conclusion: The MST-C seems useful for assessing PS performance conceptualized as the ability to discriminate complex and realistic spatial contexts. Future studies are welcome to evaluate the validity of the MST-C task as a measure of PS in clinical populations.

Keywords: Pattern separation; context discrimination; mnemonic similarity task; Mnemonic Similarity Task – Object Version; Behavioral Pattern Separation Task

Introduction

Pattern separation (PS) is the process by which similar inputs are transformed into separate non-overlapping representations. It is a critical function of episodic memory and has a notable survival value, since it allows people to compare every-day situations to those previously encountered, thus eliciting an appropriate behavioral response.¹

There are two main lines of research on PS in humans. The first line shows that impaired PS performance is a marker of mild cognitive impairment (MCI)^{2,3} and is associated with age-related specific memory impairments.^{4,5} The phenomenon seems to rely critically on poor neurogenesis at the dentate gyrus (DG) of the hippocampus.^{6,7}

A second line of research has been developed based on the observation that hippocampal neurogenesis is involved in stress, mood, and fear regulation,⁸ and that hippocampal neurogenesis acts as a buffer against stress.⁹ Thus, a deficit in PS has been hypothesized in mood or anxiety disorders.^{1,6,10} The findings currently available in the literature suggest that PS triggers excessive overgeneralization of fear or negative memories.¹¹ In anxiety or in post-traumatic stress disorders, patients tend to develop an overgeneralized fear response to neutral stimuli that resemble the traumatic experience.¹¹ Better accuracy

in recognizing similarities was observed in images encoded during periods of threat and retrieved in safety than for those encoded during periods of safety¹² or to those retrieved during a threatening situation.¹¹ In addition, in healthy subjects with subclinical depressive symptoms, lower PS performance was correlated with higher depressive symptom severity¹³ and, in mood disorder subjects, autobiographical memory retrieval was characterized by difficulty recollecting details,¹⁴ which led to “overgeneralized” memories.¹⁵ This deficit has been hypothesized as neurogenesis-dependent, since individuals with mild or moderate depressive symptoms and a low level of aerobic exercise, which is considered associated with decreased neurogenesis,¹⁶ have shown lower PS performance than subjects with less severe depressive symptoms. Moreover, impaired DG/cornu ammonis 3 (CA3) activity and poor PS performance have been observed in individuals with subclinical depressive symptoms.^{13,17}

The relationship between PS and difficulty recollecting details and recalling appropriate memories has been associated with reduced discrimination between present and past experiences in the memory, which may also result in cognitive “mistakes” in psychosis. The effect of reduced neurogenesis and/or reduced glutamate transmission observed in the DG in schizophrenia has been hypothesized to disadvantage PS, thus disadvantaging the orthogonalization of hippocampal representations of similar but distinct events. This phenomenon might advantage inappropriate association, generate false or illogical memories, and create a susceptibility to psychosis.^{18,19}

Unfortunately, the literature on PS in psychiatric disorders is still poor and lacks promising findings that can contribute to definitions of the neural correlates and mechanisms underlying these clinical features. Thus, syndromal models could be developed with multiple and complex etiologies and could include overlapping genetic and environmental risks.

PS performance has usually been measured with the Mnemonic Similarity Task – Object Version (MST-O),⁵ previously called Behavioral Pattern Separation Task – Object Version (BPS-O).²⁰ The MST-O has been used in studies evaluating age-related decline of mnemonic discrimination,^{4,5,20-28} verifying the relationship between PS and anxiety in healthy individuals,^{12,29} and assessing PS in subjects with depressive symptoms.^{13,30} This is a computerized task that tests PS performance through a recognition memory paradigm involving colored photographs of simple everyday objects on a white background. The task seems to detect the specificity of hippocampal regions involved in different types of mnemonic discrimination. Indeed, functional magnetic resonance imaging (fMRI) research has shown that regions in bilateral CA3/DG were consistent with PS measured via the BPS-O. Bakker et al.²¹ found that activity in the bilateral CA3/DG in response to the presentation of a similar BPS-O item was significantly different from the activity in response to an old item. Similarly, Yassa et al.²⁸ found a significant negative correlation between activity in the CA3/DG regions of the hippocampus and task performance, thus confirming a specific role for CA3/DG in PS. Moreover, the BPS-O was found sensitive in detecting specific age-related decline in mnemonic discrimination and in discriminating between MCI and unimpaired aging subjects.^{20,26} Unfortunately, no studies are available on the validity or reliability of the MST-O as a measure of PS.

Since it involves images of simple objects on a white background, the BPS-O does not seem fully adequate to manipulate and test context similarity,³¹ given that memories usually also include information on what is encountered in the context.³¹ In real-life, objects are presented within a rich, detailed mosaic of other features³² which creates a spatial context. Spatial context is part of the episodic memory³³ and brain areas, such as the parahippocampal place area (PPA), are involved in processing global scene information.³⁴ The fact that the MST-O does not take spatial context into account would seem an important limitation, since perceptual complexity affects short-term visual memory capacity,³⁵ which, in turn, relates to an increased perceptual similarity between complex items deriving from a large amount of overlapping information.³⁶ In addition, a complex stimulus requires greater cognitive resources than a simple one to be represented and maintained in the memory³⁶; thus, assessing subjects using simple stimuli might underestimate the performance of those with greater cognitive resources. Finally, a recent study by Libby et al.³¹ found that hippocampal neural activity changed according to item-context manipulation. In brief, paradigms using colored photographs of everyday simple objects on a white background only roughly assess PS performance, since they do not consider behavioral

context discrimination. In this framework, we propose a new task for assessing PS performance, the Mnemonic Similarity Task – Context Version (MST-C), which is based on real, complex, contextual photographs and is thus able to optimize the MST experimental paradigm in terms of the salience and generalization of the results. We here present evidence on convergent and discriminant validity regarding the association between MST-C task scores, as a measure of PS, and other variables. Convergent evidence is the degree to which test scores are correlated with tests of related constructs. Discriminant evidence is the degree to which test scores are uncorrelated with tests of unrelated constructs. Test-retest reliability, i.e., the degree to which an individual will obtain different results on a retest, was also analyzed.

Methods

Participants

Fifty healthy volunteers (50% females and 50% males; mean age: 30.66±11.73 years, age range: 21-59 years) were recruited from the general population of the Florence (Tuscany, Italy) metropolitan area. The exclusion criteria were: current or lifetime diagnosis of psychiatric disorder, assessed with the Mini International Neuropsychiatric Interview (MINI)³⁷; current MCI, i.e., a score less than 24/30 on the Mini Mental State Examination (MMSE)³⁸ or less than 26/30 at the Montreal Cognitive Assessment (MoCA)³⁹; current or lifetime neurological disorder; current use of psychotropic medications; or currently undergoing psychotherapy. These additional exclusion criteria were assessed by a supplemental set of previously used interview-based screening questions that also investigate the first-degree relatives' history of psychiatric disorders.⁴⁰ Such criteria were chosen to exclude interferences in memory and PS performance due to general cognitive impairment, drug effects, psychiatric disease, or psychological problems, as suggested in the literature.⁴¹

The majority of the sample was college students (44%), followed by white-collar workers (32%), blue-collar workers (12%), and unemployed individuals (12%). Most had a bachelor's degree (56%), 34% had a high school diploma, and 10% had a middle school diploma. The majority (88%, n= 44) had no family history of psychiatric disorders; among the remaining subjects (12%, n=6), one had a family history of panic disorder and five had a family history of major depressive episodes.

Procedures

After having provided written informed consent, the subjects were assessed by trained psychologists with the MINI, the MMSE, and the MoCA to verify eligibility. If eligible, they completed the Stroop test,⁴² the Immediate Visual Memory (IVM) test,⁴³ the MST-O,⁵ and the MST-C. To avoid familiarity and learning biases, the MST-O and the MST-C were administered in a randomized counter-balanced order. After 3 months, the participants repeated the MST-O and the MST-C.

Instruments

The MINI³⁷ is a brief structured interview extensively used as a diagnostic instrument for psychiatric disorders according to DSM-IV-TR criteria.⁴⁴ Its reliability and validity are well established⁴⁵; in the original study, kappa coefficients for concordance were good or very good for most diagnoses (from 0.53 and 0.82). The MINI's specificity was good for all diagnoses (kappa coefficients from 0.72 to 0.97), and its test-retest reliability was acceptable (kappa coefficients between 0.76 and 0.93). The Italian version showed good inter-rater validity (kappa values above 0.73) and acceptable test-retest reliability (kappa values from moderate to good agreement).⁴⁶

The Folstein MMSE³⁸ is a 30-item tool assessing cognitive functions (i.e., attention and orientation, memory, registration, recall, calculation, language, and ability to draw a complex polygon). Its total score ranges from 0 to 30, with scores of 24 or higher indicating no cognitive impairment and scores under 18 indicating severe cognitive impairment.⁴⁷ In a non-clinical community setting, the MMSE had a pooled sensitivity of 85.1% and a specificity of 85.5%.⁴⁸ The sensitivity and specificity of the Italian version were 85.7% and 90%, respectively, in a clinical sample.⁴⁹

The MoCA³⁹ is a screening instrument for MCI. It evaluates 11 cognitive domains: visuo-constructional and executive skills, naming, memory, attention, language, abstraction, delayed recall, orientation. The total maximum score is 30; a score of 26 or higher is considered as no MCI.³⁹ The MoCA was found to have very good test-retest reliability ($r = 0.92$) and internal consistency (Cronbach's $\alpha = 0.83$). The Italian version⁵⁰ showed good convergent evidence with MMSE scores ($\rho = 0.486$; $p < 0.01$),⁵¹ high sensitivity and specificity, accuracy of 0.96 (95% confidence interval), and good intra-rater, as well as inter-rater reliability with respect to the MMSE.⁵²

The Stroop Color and Word Test – Short Version⁴² assesses executive functions (i.e., selective attention, effectiveness of concentration, cognitive flexibility, processing speed, and the ability to suppress automatic responses and interference inhibition). It includes three phases: in the first and in the second phases, the subject is asked to read a series of colored words and name their color, with attention measured according to the number of mistakes. The third phase is similar to the previous phases, except the words are all names of colors; interference control is measured in terms of reaction time (i.e., interference effect in time) and the number of mistakes (i.e., interference effect in mistakes), and is obtained by subtracting the third phase performance from the sum of the first and second phase performance.⁴² The cutoff of interference measures, adjusted for age and sex, is 36.91 seconds for interference effect on reaction time and 4.23 for interference effect on mistakes; above these thresholds, the subject is considered impaired.⁴² The Stroop test has good test-retest reliability ($r > 0.80$).⁵³ The Short Version has been evaluated in a normative Italian sample ($n=248$; mean age: 52.1 ± 19.56 years, range 20-89 years).⁴²

The IVM is part of the Mental Deterioration Battery⁴³ and evaluates immediate visual memory. The subject is asked to identify, among four alternatives, figures previously

presented for a short time. The maximum score is 22; scores under 13.8, adjusted for age and education, are considered a sign of memory impairment. Single tests, as well as the entire battery, discriminate normal controls from demented patients with an accuracy of 95%.⁴³

The MST-O⁵ has a first encoding phase in which the subject judges whether 128 items are indoor or outdoor. Then, there is a recognition memory phase of 192 items in which the subject judges whether they are old, similar, or new. The results are reported as rate of correct responses for each category of items.

The Mnemonic Similarity Task – Context Version (MST-C)

We here describe the procedure followed to create the MST-C. First, 110 pairs of similar photographs of real indoor or outdoor settings were collected, along with 128 unrelated items. Similar images shared the same context (e.g., a street with cars and trees), but some details of background were different (e.g., car color or model or the location of trees) (Figure 1).

The degree of perceptual similarity of each pair of similar images was evaluated according to the literature⁵⁴ in a normative sample ($n=70$; males: 40%; mean age \pm standard deviation [SD]: 24.43 ± 2.97 years); thus, subjects were asked to express their opinion on the level of similarity of each pair of images based on a 10-point Likert scale (from 1 = different; to 10 = identical). Pairs with a score in the highest (> 75 percentile) and in the lowest (< 25 percentile) range of perceptual similarity were removed, leaving 64 pairs of similar photos (32 pairs of indoor and 32 pairs of outdoor items) and 128 unrelated items (64 indoor and 64 outdoor). This first 256-item version of the MST-C was programmed in E-Prime 2.0 software⁵⁵ and pilot tested with a sample of 20 subjects (males: 40%; mean \pm SD age: 22.75 ± 2.19 years). Based on the rate of correct responses per pair and on the suggestion that items may be best studied when whole experimental samples yield 50% correct responses,⁵⁶ 13 similar pairs correctly labeled by less than 25% of the subjects and four pairs correctly labeled by more than 75% were considered too difficult or too easy, respectively, and were deleted and replaced with new photographs. A second pilot test was run in a new sample of 20 subjects (males: 33%; mean \pm SD age = 25.25 ± 3.17 years) and, due to the rate of correct responses per pair, all items were retained. At this point, convergent evidence between MST-C and MST-O scores was preliminarily tested. The average rates of similar items correctly identified on the MST-C and MST-O were 39.9 ± 13.05 and 48.25 ± 21.24 , respectively, and the difference between the two rates (Wilcoxon test for dependent samples) was not statistically significant. The PS scores and rate of correct responses for both tasks were significantly correlated ($r = 0.487$; $p < 0.05$ and $r = 0.550$; $p < 0.05$, respectively).

The final version of the MST-C (Figure 2) includes 128 contextual images (64 indoor and 64 outdoor). In the first phase, each item is presented on a computer screen for 2 seconds followed by 0.5 seconds of interval and the subject is asked to judge whether the items are indoor or outdoor by pressing the V key or the N key, respectively.

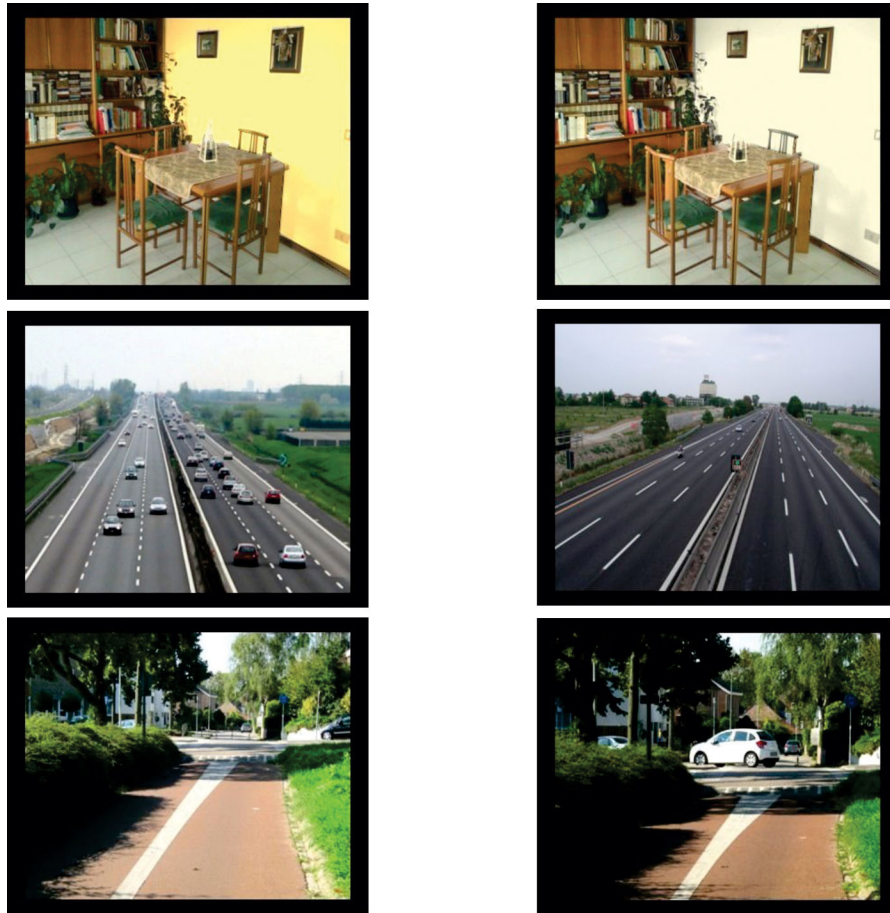


Figure 1 Example of similar items of the Mnemonic Similarity Task – Context Version (MST-C)

In the second phase (i.e., unexpected recognition memory test), the subject is asked to classify a second series of images as either old, similar, or new compared to the first series of items. Thus, 192 items are presented; 64 are repetitions of items shown in the first phase (i.e., old), 64 are new items, and 64 are altered versions of previously shown items (i.e., similar); these altered versions were obtained by modifying the original setting or by digitally altering the images using Picasa, PhotoFiltre, IrfanView and/or MS Paint software. The results are reported as rate of correct responses for each category of items.

Statistical analysis

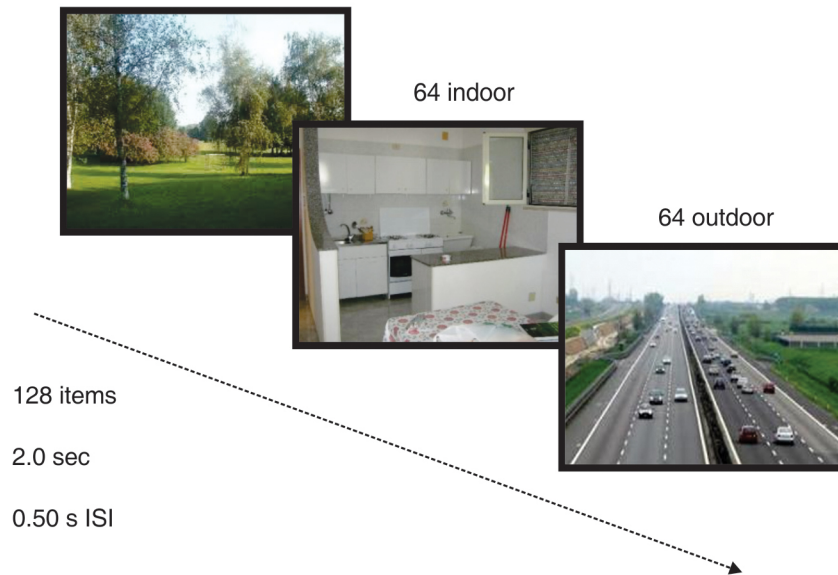
For the MST tasks, the PS score was calculated using two measures: a) the rate of similar items correctly identified minus the rate of similar items misidentified as new (SIS-SIN)⁵; b) the rate of similar items correctly identified minus the rate of similar items misidentified as old (SIS-OIS).²⁶ The recognition memory score was also calculated as the rate of old items correctly identified minus the rate of new items misidentified as old.⁵ The number of correct responses for each category of items (i.e., old, similar, new) and the type of errors (i.e., identifications of new items as similar⁵; identification of similar items as old²⁶ were also calculated.

The sample was stratified according to the following variables: sex; presence vs. absence of current or past psychiatric disorders among first-degree relatives (12% with vs. 88% without); education (median = 13 years of education; high: > 13 years of education; low: ≤ 13 years); and age (median = 26 years old; high: > 26 years old; low: ≤ 26 years old).

The influence of sociodemographic (i.e., sex, age, education, employment) and clinical variables (i.e., family history of psychiatric illness, MMSE, MoCA, Stroop test, IVM, MST-O, and MST-C) on MST performance was evaluated with the Mann Whitney *U* test for non-normally distributed variables. The Wilcoxon test for dependent samples was used to compare the rate of correct or incorrect responses on the MST-O and MST-C.

Spearman-Brown coefficients were calculated to examine: a) convergent evidence between the MST-C and MST-O; b) discriminant evidence between MST-C scores and immediate visual memory and executive functions; c) test-retest reliability of the MST-C and MST-O. The analyses were adjusted for family history for psychiatric disorders since it influenced PS in MST-O and memory in MST-C. Subjects with a family history of psychiatric disorders showed significantly lower scores in recognition memory (24.59 ± 28.77 vs. 47.29 ± 14.12 ; $p = 0.02$) and PS performance (16.18 ± 14.47 vs. 37.01 ± 20.45 ; $p = 0.01$) than the rest of the sample.

1. Encoding phase: indoor or outdoor?



2. Test phase: old, similar or new?



Figure 2 Mnemonic Similarity Task – Context Version (MST-C)

All analyses were conducted in SPSS version 20.0 for Windows. Significance levels were set at $p < 0.05$, and the strength of Spearman-Brown coefficients were evaluated according to cutoffs found in the literature⁵⁷ (i.e., r from 0.00 to 0.30 negligible correlation; r from 0.30 to 0.50 low; r from 0.50 to 0.70 moderate; r from 0.70 to 0.90 high; r from 0.90 to 1.0 very high correlation).

Results

The mean scores of MMSE, MoCA, and IVM were 29.46 ± 1.01 , 27.84 ± 1.17 , and 19.36 ± 1.18 , respectively. Stroop mistakes were 1.76 ± 0.84 , Stroop interference effect in time was 21.95 ± 7.25 seconds, and Stroop interference effect in mistakes was 1.76 ± 0.85 . Table 1 shows the

Table 1 Comparisons of MST-C and MST-O scores according to sociodemographic variables

	Gender			Education*			Age†		
	Total (n=50)	Males (n=25)	Females (n=25)	Low (n=22)	High (n=28)	p-value	Low (n=25)	High (n=25)	p-value
MST-O: PS(SIS-OIS)	6.69±33.99	14.54±37.91	-1.16±28.15	11.39±30.76	2.99±36.44	0.639	9.67±29.03	3.70±38.69	0.749
MST-O: PS(SIS-SIN)	34.51±20.84	37.87±21.74	31.15±19.75	35.86±19.58	33.45±22.07	0.222	33.90±18.60	35.11±23.23	0.742
MST-C: PS(SIS-OIS)	13.54±18.38	13.31±18.12	13.77±19.02	15.75±19.59	11.80±18.02	0.892	15.20±15.49	11.87±21.07	0.479
MST-C: PS(SIS-SIN)	17.06±15.52	15.87±14.05	18.24±17.08	18.21±14.58	16.15±16.43	0.839	16.56±12.54	17.55±18.28	0.930

Data presented as mean ± standard deviation.

MST-C = Mnemonic Similarity Task – Context Version; MST-O = Mnemonic Similarity Task – Object Version; PS = pattern separation; SIS-OIS = correct identifications as similar items minus incorrect identifications as similar items as old; SIS-SIN = correct identifications as similar items minus incorrect identifications as similar items as new.

Mann Whitney U test for independent samples.

* High education = > 13 years of education; low education = ≤ 13 years of education.

† High age = > 26 years; low age = ≤ 26 years.

means and SD of MST-C and MST-O scores for each sociodemographic variable. PS performance was not influenced by sex, age, or education level in either MST task. However, considering the effect size, even though the p-value was not statistically significant, the sex difference in MST-O score, when measured with the PS (SIS-OIS) formula, deserves consideration since it was approximately 15 points. Females performed worse than males when the PS(SIS-OIS) formula was applied to the MST-O, but not when applied to the MST-C.

Table 2 shows Spearman-Brown coefficients between PS scores in the two tasks. A statistically significant positive correlation was found between PS scores on the MST-C and the MST-O. Neither MST-C nor MST-O scores were significantly correlated with performance on the Stroop test or with IVM scores. Moreover, PS performance measured via the PS(SIS-OIS) formula in MST-C and in MST-O was not correlated with recognition memory ($r_s = -0.011$; $p = 0.930$ and $r_s = -0.118$; $p = 0.416$, respectively), while PS performance measured with the PS(SIS-SIN) formula in MST-C was correlated with recognition memory ($r_s = 0.418$; $p = 0.003$).

The rates of correct responses in MST-O and MST-C showed statistically significant correlations for new ($r_s = 0.456$; $p = 0.001$) and similar items ($r_s = 0.384$, $p = 0.006$), but not for old items ($r_s = 0.215$; $p = 0.134$).

MST-C test-retest reliability was $r_s = 0.595$ ($p = 0.000$) for the PS scores SIS-SIN; $r_s = 0.579$ ($p = 0.000$) for the PS scores SIS-OIS; and $r_s = 0.556$ ($p = 0.000$) for recognition memory score. MST-O test-retest reliability was $r_s = 0.603$ ($p = 0.000$) for the PS scores SIS-SIN; $r_s = 0.543$ ($p = 0.000$) for the PS scores SIS-OIS; and $r_s = 0.593$; $p = 0.000$ for recognition memory score.

Discussion

According to the cutoff provided in the literature,⁵⁷ the MST-C showed low convergent evidence with the MST-O, high discriminant evidence with the IVM and Stroop test, and moderate test-retest reliability. Concerning the convergent evidence, PS performance was weakly correlated in the two MST tasks. This result is consistent with the hypothesis that PS performance conceptualized as real-life context discrimination and PS performance conceptualized as simple object discrimination do not overlap and that MST-O is inadequate for measuring PS performance in context.

Concerning discriminant evidence, there was no correlation between MST-C PS performance and attention, interference control, or immediate visual memory. This result is consistent with Toner et al.²⁶ who found no correlation between the Stroop test and rates of correct identifications of similar MST-O items, as well as with Stark et al.,⁵⁸ who found no statistically significant correlations between PS performance and digit span performance or intelligence quotient. Thus, PS performance seems independent of executive function and memory.²⁶

MST-C PS score and recognition memory score were not correlated when the PS(SIS-OIS) formula was applied, and they were poorly correlated when the PS(SIS-SIN) formula was applied. These findings are consistent with

Table 2 Spearman-Brown coefficients between MST-C and MST-O, Stroop Color and Word Test, IVM

	MST-O: PS (SIS-SIN)	MST-C: PS (SIS-SIN)	Stroop test: attention	Stroop test: time	Stroop test: mistakes	IVM
MST-O: PS (SIS-SIN)	0.60	0.77	-0.13	-0.17	-0.05	N/A
MST-C: PS (SIS-SIN)	0.46	0.59	-0.38	0.08	-0.13	N/A
Stroop test: attention	-0.08	-0.24	0.68 ⁵³	0.21	0.08	N/A
Stroop test: time	-0.09	0.04	0.12	0.48 ⁵³	0.85	N/A
Stroop test: mistakes	-0.03	-0.07	0.05	0.43	0.53 ⁵³	N/A
IVM	0.14	0.01	-0.15	-0.42	-0.34	N/A

Test-retest reliability is reported on the diagonal. Raw Spearman-Brown correlations are presented below the diagonal. Correlations disattenuated for unreliability of both the x and y variables are presented above the diagonal.

IVM = Immediate Visual Memory; MST-C = Mnemonic Similarity Task – Context Version; MST-O = Mnemonic Similarity Task – Object Version; N/A = not available; PS = pattern separation; SIS-SIN: the number of correctly identified similar items minus the number of new items misidentified as similar ones.

the literature.^{4,5,26,59} Thus, the ability to distinguish among similar items seems unrelated to memory performance.

Concerning test-retest reliability, MST-C PS performance was highly stable after three months. Although the literature suggests that age influences PS performance,^{4,27} it did not in the present study. This inconsistency might be explained by the relatively young mean age of our sample.

The effect size observed for sex difference in the MST-O should be further explored. It has been recently suggested that, in rats, there are sex differences in neurogenesis regulation and PS, as well as in hippocampal function and neural plasticity.⁶⁰ However, this result should also be evaluated in light of the fact that the effect size disappeared when PS was measured with the MST-C. Thus, again, spatial context might play a key role in PS performance assessment.

Some limitations should be pointed out. First, the sample may seem small and not fully representative of the general population, since subjects were recruited by convenience sampling. However, a number of studies on computerized cognitive tasks have been conducted with samples of less than 50 individuals,⁶¹⁻⁶³ and convenience sampling is a standard procedure in this kind of research.⁵⁴ Second, although we measured the degree of perceptual and mnemonic similarity of each pair of similar stimuli, we did not provide a corresponding gradient in behavioral output, as recommended in recent publications.⁴¹ However, the results of the current study may be used in the future to provide a mnemonic similarity gradient for similar MST-C items. Third, the potentially high overall degree of semantic similarity among MST-C items might have biased the response. It could be claimed that to validly assess behavioral PS performance, novel abstract images would be preferable to familiar items in order to correct for semantic effects. However, activation in hippocampal regions has been observed^{21,25,28,41,64} in studies employing familiar items and in those using abstract, novel items.^{65,66} In addition, this potential bias was overcome by considering MST performance as the rate of correct responses to a similar item minus the rate of similar responses given to a new item.²⁰

Despite these limitations, the MST-C can surpass current instruments since it is a measure of PS performance that considers spatial context and has known psychometric properties: low convergent evidence, high discriminant

evidence, moderate test-retest reliability. Thus, the MST-C can be considered useful for assessing PS performance in healthy subjects in real-life contexts and might be useful in clinical settings as a proxy for episodic memory function. From a research point of view, the MST-C may facilitate the study of neural correlates of PS performance in neuroimaging research, as well as help provide insight into the underlying etiopathogenetic mechanisms of psychiatric disorders. It could also be helpful in clinical trials, once the learning effects are controlled, due to its good test-retest reliability. Having a prevalence of visual content, it is easy to translate and adapt into other cultures, thus multicentric investigations and transcultural research could be undertaken.

Studies of healthy older adults, clinical populations (e.g., patients with cognitive impairment or affective, psychotic, or obsessive-compulsive-related disorders) or other variables are needed to confirm these preliminary findings and determine whether PS performance, as assessed by the MST-C, is influenced by semantic similarities or by an inter-item perceptual similarity effect, as well as to test the validity evidence of the MST-C as a measure of memory.

Disclosure

The authors report no conflicts of interest.

References

- 1 Kheirbek MA, Klemenhagen KC, Sahay A, Hen R. Neurogenesis and generalization: a new approach to stratify and treat anxiety disorders. *Nat Neurosci.* 2012;15:1613-20.
- 2 Bakker A, Albert MS, Krauss G, Speck CL, Gallagher M. Response of the medial temporal lobe network in amnesic mild cognitive impairment to therapeutic intervention assessed by fMRI and memory task performance. *Neuroimage Clin.* 2015;7:688-98.
- 3 Yassa MA, Stark SM, Bakker A, Albert MS, Gallagher M, Stark CE. High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in patients with amnesic mild cognitive impairment. *Neuroimage.* 2010;51:1242-52.
- 4 Holden HM, Toner C, Pirogovsky E, Kirwan CB, Gilbert PE. Visual object pattern separation varies in older adults. *Learn Mem.* 2013;20:358-62.
- 5 Stark SM, Stevenson R, Wu C, Rutledge S, Stark CE. Stability of age-related deficits in the mnemonic similarity task across task variations. *Behav Neurosci.* 2015;129:257-68.
- 6 Clemenson GD, Stark CE. Virtual environmental enrichment through video games improves hippocampal-associated memory. *J Neurosci.* 2015;35:16116-25.

- 7 França TF, Bitencourt AM, Maximilla NR, Barros DM, Monserrat JM. Hippocampal neurogenesis and pattern separation: a meta-analysis of behavioral data. *Hippocampus*. 2017;27:937-50.
- 8 Anacker C. Adult hippocampal neurogenesis in depression: Behavioral implications and regulation by the stress system. *Curr Top Behav Neurosci*. 2014;18:25-43.
- 9 Snyder JS, Soumier A, Brewer M, Pickel J, Cameron HA. Adult hippocampal neurogenesis buffers stress responses and depressive behaviour. *Nature*. 2011;476:458-61.
- 10 Samuels BA, Hen R. Neurogenesis and affective disorders. *Eur J Neurosci*. 2011;33:1152-9.
- 11 Balderston NL, Mathur A, Adu-Brimpong J, Hale EA, Ernst M, Grillon C. Effect of anxiety on behavioural pattern separation in humans. *Cogn Emot*. 2017;31:238-48.
- 12 Segal SK, Stark SM, Kattan D, Stark CE, Yassa MA. Norepinephrine-mediated emotional arousal facilitates subsequent pattern separation. *Neurobiol Learn Mem*. 2012;97:465-9.
- 13 Shelton DJ, Kirwan CB. A possible negative influence of depression on the ability to overcome memory interference. *Behav Brain Res*. 2013;256:20-6.
- 14 Lemogne C, Piolino P, Friszer S, Claret A, Girault N, Jouvent R, et al. Episodic autobiographical memory in depression: specificity, autonoetic consciousness and self-perspective. *Conscious Cogn*. 2006;15:258-68.
- 15 Sperduti M, Martinelli P, Kalenzaga S, Devauchelle AD, Lion S, Malherbe C, et al. Don't be too strict with yourself! rigid negative self-representation in healthy subjects mimics the neurocognitive profile of depression for autobiographical memory. *Front Behav Neurosci*. 2013;7:41.
- 16 Fabel K, Wolf SA, Ehninger D, Babu H, Leal-Galicia P, Kempermann G. Additive effects of physical exercise and environmental enrichment on adult hippocampal neurogenesis in mice. *Front Neurosci*. 2009;3:50.
- 17 Fujii T, Saito DN, Yanaka HT, Kosaka H, Okazawa H. Depressive mood modulates the anterior lateral CA1 and DG/CA3 during a pattern separation task in cognitively intact individuals: a functional MRI study. *Hippocampus*. 2014;24:214-24.
- 18 Tamminga CA, Stan AD, Wagner AD. The hippocampal formation in schizophrenia. *Am J Psychiatry*. 2010;167:1178-93.
- 19 Das T, Ivleva EI, Wagner AD, Stark CE, Tamminga CA. Loss of pattern separation performance in schizophrenia suggests dentate gyrus dysfunction. *Schizophr Res*. 2014;159:193-7.
- 20 Stark SM, Yassa MA, Lacy JW, Stark CE. A task to assess behavioural pattern separation (BPS) in humans: data from healthy aging and mild cognitive impairment. *Neuropsychologia*. 2013;51:2442-9.
- 21 Bakker A, Kirwan CB, Miller M, Stark CE. Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science*. 2008;319:1640-2.
- 22 Holden HM, Hoebel C, Loftis K, Gilbert PE. Spatial pattern separation in cognitively normal young and older adults. *Hippocampus*. 2012;22:1826-32.
- 23 Kirwan CB, Stark CE. Overcoming interference: an fMRI investigation of pattern separation in the medial temporal lobe. *Learn Mem*. 2007;14:625-33.
- 24 Kheirbek MA, Hen R. Add neurons, subtract anxiety. *Sci Am*. 2014;311:62-7.
- 25 Lacy JW, Yassa MA, Stark SM, Muftuler LT, Stark CE. Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learn Mem*. 2011;18:15-8.
- 26 Toner CK, Pirogovsky E, Kirwan CB, Gilbert PE. Visual object pattern separation deficits in nondemented older adults. *Learn Mem*. 2009;16:338-42.
- 27 Yassa MA, Mattfeld AT, Stark SM, Stark CE. Age-related memory deficits linked to circuit-specific disruption in the hippocampus. *Proc Natl Acad Sci U S A*. 2011;108:8873-8.
- 28 Yassa MA, Lacy JW, Stark SM, Albert MS, Gallagher M, Stark CE. Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus*. 2011;21:968-79.
- 29 Balderston NL, Mathur A, Adu-Brimpong J, Hale EA, Ernst M, Grillon C. Effect of anxiety on behavioural pattern separation in humans. *Cogn Emot*. 2015;31:238-48.
- 30 Déry N, Pilgrim M, Gibala M, Gillen J, Wojtowicz JM, Macqueen G, et al. Adult hippocampal neurogenesis reduces memory interference in humans: opposing effects of aerobic exercise and depression. *Front Neurosci*. 2013;7:66.
- 31 Libby LA, Ragland JD, Ranganath C. The hippocampus generalizes across memories that share item and context information[Internet]. 2016 Apr 22[cited 2017 Sep 29]. www.biorxiv.org/content/early/2016/04/22/049965.
- 32 Oliva A, Torralba A. The role of context in object recognition. *Trends Cogn Sci*. 2007;11:520-7.
- 33 Rolls ET. Pattern separation, completion, and categorisation in the hippocampus and neocortex. *Neurobiol Learn Mem*. 2016;129:4-28.
- 34 Brockmole JR, Henderson JM. Using real-world scenes as contextual cues for search. *Vis Cogn*. 2006;13:99-108.
- 35 Luria R, Sessa P, Gotler A, Jolicœur P, Dell'Acqua R. Visual short-term memory capacity for simple and complex objects. *J Cogn Neurosci*. 2010;22:496-512.
- 36 Jackson MC, Linden DE, Roberts MV, Kriegeskorte N, Haenschel C. Similarity, not complexity, determines visual working memory performance. *J Exp Psychol Learn Mem Cogn*. 2015;41:1884-92.
- 37 Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59:22-33.
- 38 Folstein MF, Folstein SE, McHugh PR. "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *J Psychiatry Res*. 1975;12:189-98.
- 39 Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53:695-9.
- 40 Cosci F, Abrams K, Schruers KR, Rickelt J, Griez EJ. Effect of nicotine on 35% CO2 induced anxiety: a study in healthy volunteers. *Nicotine Tob Res*. 2006;8:511-7.
- 41 Liu KY, Gould RL, Coulson MC, Ward EV, Howard RJ. Tests of pattern separation and pattern completion in humans -- a systematic review. *Hippocampus*. 2016;26:705-17.
- 42 Caffarra P, Vezzadini G, Dieci F, Zonato F, Venneri A. Una versione abbreviata del test di Stroop: dati normativi nella popolazione italiana. *Riv Neurol*. 2002;12:111-5.
- 43 Carlesimo GA, Caltagirone C, Gainotti G. The mental deterioration battery: normative data, diagnostic reliability and qualitative analyses of cognitive impairment. The Group for the Standardization of the Mental Deterioration Battery. *Eur Neurol*. 1996;36:378-84.
- 44 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)*. Arlington: American Psychiatric Publishing; 2000.
- 45 Lecrubier Y, Sheehan DV, Weiller E, Amorim P, Bonora I, Harnett Sheehan K, et al. The Mini International Neuropsychiatric Interview (M.I.N.I.). A short diagnostic structured interview: reliability and validity according to the CIDI. *Eur Psychiatry*. 1997;12:224-31.
- 46 Rossi A, Alberio R, Porta A, Sandri M, Tansella M, Amaddeo F. The reliability of the mini-international neuropsychiatric interview – Italian version. *J Clin Psychopharmacol*. 2004;24:561-3.
- 47 Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. *J Am Geriatr Soc*. 1992;40:922-35.
- 48 Mitchell AJ. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *J Psychiatr Res*. 2009;43:411-31.
- 49 Cossa FM, Della Sala S, Musicco M, Spinnler H, Ubezio MC. Comparison of two scoring systems of the mini-mental state examination as a screening test for dementia. *J Clin Epidemiol*. 1997;50:961-5.
- 50 Pirani A, Nasreddine ZS, Tulipani C, Zaccherini D, Martini E, Neri M. Montreal Cognitive Assessment (MoCA): uno strumento rapido per lo screening del mild cognitive impairment. Dati preliminari della versione Italiana. In: *Atti IV Congresso Regionale Associazione Italiana Psicogeriatrica*; 2007; Bologna, Italia.
- 51 Conti S, Bonazzi S, Laiacoma M, Masina M, Coralli MV. Montreal Cognitive Assessment (MoCA)-Italian version: regression based norms and equivalent scores. *Neurol Sci*. 2015;36:209-14.
- 52 Pirrotta F, Timpano F, Bonanno L, Nunnari D, Marino S, Bramanti P, et al. Italian validation of Montreal cognitive assessment. *Eur J Psychol Assess*. 2015;31:131-7.

- 53 Siegrist M. Test-retest reliability of different versions of the Stroop test. *J Psychol.* 1997;131:299-306.
- 54 Kirwan CB, Stark CE. Overcoming interference: an fMRI investigation of pattern separation in the medial temporal lobe. *Learn Mem.* 2007;14:625-33.
- 55 Schneider W, Eschman A, Zuccolotto A. E-prime user's guide-Pittsburgh: Psychology Software Tools; 2012.
- 56 Kelley TL. The selection of upper and lower groups for the validation of test items. *J Educ Psychol.* 1939;30:17-24.
- 57 Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012;24:69-71.
- 58 Stark SM, Yassa MA, Stark CE. Individual differences in spatial pattern separation performance associated with healthy aging in humans. *Learn Mem.* 2010;17:284-8.
- 59 Brock Kirwan C, Hartshorn A, Stark SM, Goodrich-Hunsaker NJ, Hopkins RO, Stark CE. Pattern separation deficits following damage to the hippocampus. *Neuropsychologia.* 2012;50:2408-14.
- 60 Yagi S, Chow C, Lieblich SE, Galea LA. Sex and strategy use matters for pattern separation, adult neurogenesis, and immediate early gene expression in the hippocampus. *Hippocampus.* 2016;26:87-101.
- 61 Leyk D, Sievert A, Heiss A, Gorges W, Ridder D, Alexander T, et al. Validation of a short-term memory test for the recognition of people and faces. *Ergonomics.* 2008;51:1125-36.
- 62 Pilli R, Naidu M, Pingali U, Shobha J. Evaluation of a new computerized psychometric test battery: effects of zolpidem and caffeine. *J Pharmacol Pharmacother.* 2013;4:247-55.
- 63 Sánchez-Cubillo I, Periañez JA, Adrover-Roig D, Rodríguez-Sánchez JM, Ríos-Lago M, Tirapu J, et al. Construct validity of the trail making test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *J Int Neuropsychol Soc.* 2009;15:438-50.
- 64 Kirwan CB, Jones CK, Miller MI, Stark CE. High-resolution fMRI investigation of the medial temporal lobe. *Hum Brain Mapp.* 2007 28:959-66.
- 65 Lissek S, Bradford DE, Alvarez RP, Burton P, Espensen-Sturges T, Reynolds RC, et al. Neural substrates of classically conditioned fear-generalization in humans: a parametric fMRI study. *Soc Cogn Affect Neurosci.* 2014;9:1134-42.
- 66 Paleja M, Girard TA, Herdman KA, Christensen BK. Two distinct neural networks functionally connected to the human hippocampus during pattern separation tasks. *Brain Cogn.* 2014;92C:101-11.