Dement Neurocogn Disord. 2025 Oct;24(4):286-300
https://doi.org/10.12779/dnd.2025.24.4.286
pISSN 1738-1495·eISSN 2384-0757

DND Dementia and Neurocognitive Disorder

Original Article

Check for updates

Clinical Utility and Diagnostic Accuracy of the Tablet-Based Seoul Cognitive Status Test: Evidence for Scalable Cognitive Assessment

Young Ju Kim [1,*] **Joon Soo Shin** [1,*] **Dayeong An** [1,2] **Duk L. Na** [1,2]
Hee Jin Kim [3,4,5,6]

[1]Beaubrain Healthcare Co., Ltd., Seoul, Korea
[2]Department of Neurology, Happymind Clinic, Seoul, Korea
[3]Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea
[4]Alzheimer's Disease Convergence Research Center, Samsung Medical Center, Seoul, Korea
[5]Neuroscience Center, Samsung Medical Center, Seoul, Korea
[6]Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences & Technology (SAIHST), Sungkyunkwan University, Seoul, Korea

Received: Sep 17, 2025
Revised: Oct 9, 2025
Accepted: Oct 17, 2025
Published online: Oct 23, 2025

Correspondence to
Hee Jin Kim
Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Korea.
Email: evekhj@gmail.com

*Young Ju Kim and Joon Soo Shin contributed equally as first authors.

ORCID iDs
Young Ju Kim
https://orcid.org/0000-0003-2927-6724
Joon Soo Shin
https://orcid.org/0000-0002-4641-6103
Dayeong An
https://orcid.org/0009-0004-1176-7597
Duk L. Na
https://orcid.org/0000-0002-0098-7592

ABSTRACT

Background and Purpose: The Seoul Cognitive Status Test (SCST) is a brief (~30 minutes), tablet-based battery that yields domain-specific and composite scores and uses automated scoring for rapid result generation, reduced examiner burden, and scalability for large or longitudinal assessments. We evaluated the SCST's psychometric and diagnostic performance against multidisciplinary final clinical diagnoses derived from a multidimensional workup.
Methods: We conducted a cross-sectional diagnostic study of 777 outpatients (SCST–Seoul Neuropsychological Screening Battery [SNSB] n=639; SCST–Consortium to Establish a Registry for Alzheimer's Disease [CERAD] n=138) who underwent standardized diagnostic evaluation; multidisciplinary final diagnoses served as the reference. Participants completed the SCST and either the Korean version of CERAD (CERAD-K) or SNSB-second edition (SNSB-II). We assessed convergent validity of SCST subtests with traditional measures and compared SCST composite scores and classifications with reference diagnoses.
Results: Pearson correlations between SCST subtests and corresponding CERAD-K/SNSB-II measures were all significant ($p<0.001$), with key domain correlations ≥0.7 (language, memory, executive). In the pooled sample, covariate-adjusted receiver operating characteristic area under the curves (AUCs) for the SCST composite ranged from 0.854 (cognitively unimpaired [CU] vs. mild cognitive impairment [MCI]) to 0.980 (CU vs. dementia); AUC for CU vs. cognitively impaired [CI] was 0.903. SCST-based classification sensitivity was 0.859 (CU vs. CI), 0.984 (CU vs. dementia), 0.758 (CU vs. MCI), 0.896 (MCI vs. dementia), and 0.882 (no dementia vs. dementia).
Conclusions: These findings support the SCST's clinical utility; further studies in broader community-based cohorts are warranted to confirm generalizability.

Keywords: Neuropsychological Tests; Cognitive Dysfunction; Dementia; Diagnosis; Computers

Hee Jin Kim
https://orcid.org/0000-0002-3186-9441

Conflict of Interest
Duk L. Na is scientific co-founders of BeauBrain Healthcare, a company aiming to commercialize the Seoul Cognitive Status Test. Young Ju Kim, Joon Soo Shin, and Dayeong An are employed at BeauBrain Healthcare. Duk L. Na and Joon Soo Shin had no role in the analysis and interpretation of the data. The remaining authors have no conflicts of interest to declare.

Author Contributions
Conceptualization: Na DL; Data curation: Shin JS; Formal analysis: Kim YJ, An D; Investigation: Shin JS, An D; Methodology: Kim YJ; Supervision: Na DL, Kim HJ; Validation: Kim YJ; Writing - original draft: Shin JS, Kim YJ; Writing - review & editing: Kim HJ.

INTRODUCTION

Dementia poses a growing global public-health burden, and timely detection of cognitive decline is critical for patient care and health-system planning. The Seoul Neuropsychological Screening Battery-second edition (SNSB-II) is a comprehensive clinician-administered battery that provides detailed domain-level profiles across attention, language, visuospatial function, memory, and executive function; the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease (CERAD-K) is a widely used standardized screening and staging battery that emphasizes word-list memory, naming, and constructional praxis for dementia evaluation.[1,2] However, these instruments are time-consuming, require trained personnel for standardized administration and scoring, can induce fatigue in older adults, and are relatively limited in scalability for repeated or large-sample assessments; consequently, their routine use in high-throughput clinical workflows is constrained.

Recent work has demonstrated that brief, tablet-based and AI-assisted digital cognitive assessments can deliver reliable domain-level measures and show promise for diagnostic and prognostic use. For example, recent studies have linked brief digital batteries and AI-enabled drawing tasks to amyloid/tau biomarkers and future cognitive decline, and have shown utility when combined with blood-based biomarkers.[3-7]

The Seoul Cognitive Status Test (SCST) was developed to address these practical limitations. The SCST is a brief (30-minute), tablet-based battery that assesses five core domains—attention, language, visuospatial function, memory, and executive function—and yields both domain and composite scores. The SCST normative sample (n=1,006) was recruited nationwide from diverse settings, including community volunteers, primary-care clinics, and tertiary memory clinics, ensuring broad geographic and clinical representativeness. In addition to shorter administration time (reducing participant fatigue), the SCST captures rich process-level metrics (item-level response times, sequence and error patterns) that enable fine-grained process analysis; provides continuous, education-adjusted normative scoring rather than coarse categorical thresholds, enhancing sensitivity to subthreshold change; and supports examiner-assisted digital administration that reduces inter-rater variability and facilitates scalable deployment for repeated measurements, remote/hybrid workflows, and integration with biomarker or trial-selection pipelines.[8,9]

Prior validation studies have reported moderate-to-strong correlations between SCST subtests and traditional paper-and-pencil measures and have shown promising discriminative performance for classifying cognitively unimpaired (CU), mild cognitive impairment (MCI), and dementia.[8,10-12] Nonetheless, some earlier studies relied primarily on comparisons with other neuropsychological tests rather than evaluation against multidisciplinary final clinical diagnoses established by experienced clinicians, limiting inference about real-world clinical utility.[9,13]

Accordingly, the present study evaluated the SCST in a newly recruited, large sample with three primary aims: 1) to assess convergent validity between SCST subtests and corresponding measures from CERAD-K and SNSB-II; 2) to evaluate the diagnostic utility of the SCST composite score against multidisciplinary final clinical diagnoses (reference standard); and 3) to assess the agreement of SCST-based categorical classification with those multidisciplinary final clinical diagnoses. By addressing both psychometric properties and clinical performance, we aimed to provide robust evidence for the SCST's applicability in routine diagnostic workflows.

METHODS

Participants

This cross-sectional diagnostic study included 777 participants from two head-to-head cohorts: SCST–CERAD (n=138) and SCST–SNSB (n=639). Most participants were recruited from a primary-care clinic (Happymind Clinic; n=731, 94.1%), with a minority (Samsung Medical Center; n=46, 5.9%) enrolled from a tertiary referral center in South Korea. All participants underwent a standardized diagnostic workup comprising structured clinical interviews, neurological examinations, comprehensive neuropsychological testing (SNSB-II or CERAD-K as appropriate), caregiver-reported functional assessment (Korean instrumental activities of daily living, K-IADL),[14] laboratory testing to exclude reversible causes of cognitive impairment, and brain imaging (magnetic resonance imaging and/or amyloid positron emission tomography) when available. Final clinical diagnoses were established through outpatient follow-up and multidisciplinary synthesis of all available data and served as the reference standard. Based on the final clinical diagnosis, participants were categorized as CU (n=187), MCI (n=355), or dementia (n=235). CU participants had no history of neurological or psychiatric disorders and exhibited normal cognitive performance on neuropsychological assessments. MCI was diagnosed based on Petersen's criteria, which require subjective cognitive complaints (reported by the patient or caregiver), objective cognitive deficits, preserved functional independence in daily living, and no evidence of dementia.[15] Dementia of the Alzheimer type was diagnosed according to the National Institute on Aging–Alzheimer's Association research criteria for probable Alzheimer's disease (AD).[16]

Participants were eligible if they completed the standardized diagnostic evaluation and received a multidisciplinary final clinical diagnosis (CU, MCI, or probable AD dementia). Exclusion criteria included severe medical conditions likely to confound cognition (operationalized using criteria adapted from Christensen et al.[17]), major untreated psychiatric disorders, and acute confusional states; stable, treated medical or psychiatric comorbidities were permitted and recorded. Low educational attainment, including illiteracy, was not an exclusion criterion: the SCST normative sample (n=1,006) spans a broad education range, SCST scores were referenced to age/sex/education-adjusted norms, and years of education were included as a covariate in adjusted analyses.

Index test—SCST

The SCST is a tablet-based cognitive battery administered on a 12-inch Android device. It comprises seven subtests across five cognitive domains, including the forward and backward tests of the Visual Span Test (VST) for attention; the Difficult Naming Test (DNT) and semantic and phonemic verbal fluency tasks for language; the Block Design Test (BDT) for visuospatial function; time orientation and the Word–Place Association Test (WPAT) for memory; and the Korean Trail Making Test–Elderly version (K-TMT-E), parts A and B, for executive function. The SCST produces domain scores and a composite score. A total of 1,006 CU older adults in South Korea were included in the SCST normative sample. To ensure representativeness, normative participants were recruited from multiple geographic regions and diverse settings (primary/secondary care clinics, tertiary hospitals, and community centers).[8,9] Administration procedures, examiner training, scoring algorithms, and the methods used to derive normative reference values (including initial standardization and subsequent re-standardization) are described in the **Supplementary Data 1** and **2**.

The SCST was examiner-assisted and administered by trained examiners according to a standardized protocol. Mean completion time was approximately 30 minutes, and examiners provided brief technical assistance for tablet operation when required. No participants in the diagnostic cohorts were excluded for inability to complete the SCST due to low digital literacy.

Comparator neuropsychological batteries

Traditional paper-and-pencil batteries used as part of the diagnostic workup included the SNSB-II and the CERAD-K. The SNSB-II assesses the Digit Span Test (DST), the Korean version of the Boston Naming Test (K-BNT), the Controlled Oral Word Association Test, the Rey-Osterrieth Complex Figure Test (RCFT) copy, the Seoul Verbal Learning Test-Elderly's version (immediate/delayed/recognition), and the K-TMT-E.[1] The CERAD-K comprises the K-BNT, semantic fluency, constructional praxis, word list memory/recall/recognition, and TMT parts A/B.[2] Corresponding subtests were pre-specified for convergent validity analyses with SCST subtests.

Reference standard—final clinical diagnosis

The final clinical diagnosis served as the reference standard (standard of truth) and was established by a multidisciplinary synthesis of all available information (comparator test results, clinical interview, neurological examination, K-IADL, laboratory studies, and imaging). Dementia severity was assessed using the Clinical Dementia Rating (CDR), where applicable.[17] Final diagnoses were assigned by experienced neurologists (D.L.N., and S.W.S., each >20 years' experience) in conjunction with the multidisciplinary team; SCST results were not made available to adjudicators and were not incorporated into the diagnostic process to avoid incorporation bias.

The majority of participating centers routinely administer the SNSB-II as part of standard diagnostic evaluation (yielding the larger SCST–SNSB cohort), while CERAD-K data were obtained at a subset of sites for complementary validation (SCST–CERAD).

SCST-based classification, test administration, and blinding

For SCST-based classification analyses, personal identifiers were removed from SCST output. De-identified SCST profiles were reviewed by two experienced clinical neuropsychologists who were blinded to the final clinical diagnoses and to the detailed comparator battery results. Based solely on SCST performance, participants were first classified as CU or cognitively impaired (CI). Participants classified as CI were further subtyped into MCI or dementia using limited additional functional information (K-IADL and CDR) provided with the de-identified profiles.[14,18,19] All assessments (SCST and comparator batteries) were administered within a predefined clinical window under standardized conditions by trained examiners. Any discrepancies in SCST-based classification were resolved by consensus between two blinded neuropsychologists (Y.J.K., and D.Y.A., each >10 years' experience).

Statistical analyses

Independent-samples *t*-tests and $tests were used to compare the demographic characteristics of the two cohorts (SCST-CERAD and SCST-SNSB).

Convergent and discriminant validity were evaluated using Pearson correlation coefficients between SCST subtest raw scores and the corresponding CERAD-K or SNSB-II scores within each cohort. For visualization, scatterplots were generated using z-scores. To enhance plot clarity, extreme values were omitted from figures (**Supplementary Data 1** for outlier rules).

https://doi.org/10.12779/dnd.2025.24.4.286

Correlation strength was interpreted according to Mukaka[20]: <0.3 negligible, 0.3–0.5 low, 0.5–0.7 moderate, 0.7–0.9 high, >0.9 very high.

As a secondary known-groups analysis, we compared SCST subtest, domain, and composite scores across CU, MCI, and dementia using analysis of covariance adjusted for age, sex, and education. Bonferroni-adjusted pairwise comparisons are reported in the **Supplementary Data 1**.

Next, we assessed the diagnostic performance of the SCST composite score using receiver operating characteristic (ROC) analyses adjusted for age, sex, and education. Reported metrics included the area under the curve (AUC), as well as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and positive likelihood ratio (PLR).

Before conducting the accuracy analysis of the SCST-based categorical classifications (CU, MCI, and dementia), we quantified the inter-rater agreement between the two neuropsychologists using Cohen's kappa.[21] Discrepant ratings were adjudicated to derive the final SCST-based categorical classifications used for the analyses. Sensitivity, specificity, PPV, NPV, false positive rate (FPR), false negative rate (FNR), and PLR were computed for these classifications by direct comparison with the reference standard.

All statistical analyses were performed using SPSS version 24.0 (IBM Corp., Armonk, NY, USA) and R version 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria). A two-sided *p*-value of less than 0.05 was considered statistically significant.

RESULTS

Characteristics of the study participants

Participant characteristics are summarized in **Table 1**. In the pooled sample (n=777) mean age was 73.7±7.6 years, mean education 13.0±4.1 years, female 485 (62.4%), and apolipoprotein E (APOE) ε4 carriers 337 (43.4%); diagnostic distribution: CU, 187 (24.1%); MCI, 355 (45.7%); dementia, 235 (30.2%). In the SCST–CERAD (n=138) and SCST–SNSB (n=639) cohorts, participants were aged 73.3±7.4 and 73.8±7.7 years, respectively, with mean education of 13.4±3.8 and 12.9 ± 4.2 years; female participants numbered 91 (65.9%) and 394 (61.7%), and APOE ε4 carriers numbered 47 (34.1%) and 290 (45.4%), respectively. The proportions of CU/MCI/dementia were 38.4%/29.7%/31.9% in the SCST–CERAD cohort and 21.0%/49.1%/29.9% in the SCST–SNSB cohort.

Table 1. Baseline demographic and clinical characteristics of the study participants

Characteristics	Participant			p-value
	Pooled sample (n=777)	SCST-CERAD cohort (n=138)	SCST-SNSB cohort (n=639)	
Age (yrs)	73.7±7.6	73.3±7.4	73.8±7.7	0.463
Female	485 (62.4)	91 (65.9)	394 (61.7)	0.346
Education (yrs)	13.0±4.1	13.4±3.8	12.9±4.2	0.213
APOE ε4 carriers	337 (43.4)	47 (34.1)	290 (45.4)	0.025
CU/MCI/dementia	187/355/235 (24.1/45.7/30.2)	53/41/44 (38.4/29.7/31.9)	134/314/191 (21.0/49.1/29.9)	<0.001

Values are presented as mean ± standard deviation or number (%).
SCST: Seoul Cognitive Status Test, CERAD: Consortium to Establish a Registry for Alzheimer's Disease, SNSB: Seoul Neuropsychological Screening Battery, APOE: apolipoprotein E, CU: cognitively unimpaired, MCI: mild cognitive impairment.

Correlations between SCST and traditional tests

Pearson correlations (all $p<0.001$) indicated strong concordance between SCST subtests and corresponding CERAD-K/SNSB-II measures, particularly in language, memory, and executive domains. Specifically, the DNT correlated highly with the K-BNT ($r=0.814$ with CERAD-K; $r=0.816$ with SNSB-II), and WPAT immediate and delayed recall correlated strongly with CERAD word list (CERAD-K; $r=0.775$ and 0.823) and with SVLT (SNSB-II; $r=0.716$ and 0.780). Executive measures also showed large effects (K-TMT-E part A vs. CERAD-K TMT-A, $r=0.886$; K-TMT-E part B vs. SNSB-II TMT-B, $r=0.790$). Visuospatial performance on the BDT demonstrated moderate correlations with CERAD-K constructional praxis ($r=0.537$) and with RCFT copy (SNSB-II; $r=0.547$). Verbal fluency on the SCST showed moderate associations with CERAD-K semantic fluency ($r=0.698$) and with SNSB-II fluency (semantic, $r=0.650$; phonemic, $r=0.683$). Recognition memory correlations were modest (WPAT recognition vs. CERAD word list recognition, $r=0.619$ [CERAD-K]; vs. SVLT recognition, $r=0.528$ [SNSB-II]), while attention correlations were lower (VST vs. DST forward, $r=0.344$; backward, $r=0.454$ [SNSB-II]). Overall, the SCST composite score correlated strongly with the CERAD-K total score ($r=0.923$) (**Fig. 1**, **Supplementary Tables 1** and **2**).

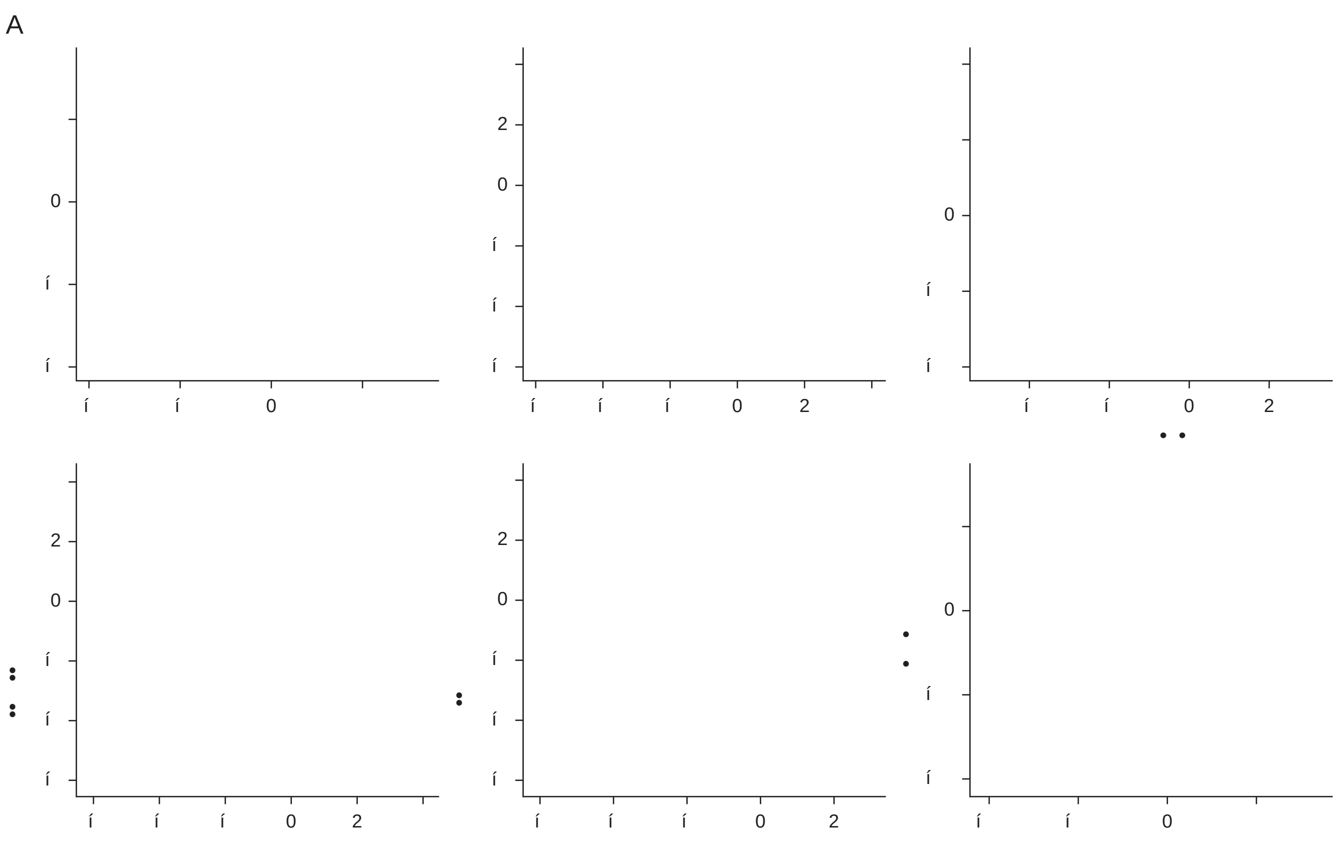

Fig. 1. Correlation scatterplots between the SCST subtests (y-axis) and corresponding subtests of the CERAD-K and SNSB-II (x-axis). (A) SCST-CERAD cohort, (B) SCST-SNSB cohort. Scatterplots are based on standardized scores for visualization.
SCST: Seoul Cognitive Status Test, CERAD-K: Korean version of CERAD, SNSB-II: second edition of SNSB, CERAD: Consortium to Establish a Registry for Alzheimer's Disease, SNSB: Seoul Neuropsychological Screening Battery, DNT: Difficult Naming Test, K-BNT: Korean version of the Boston Naming Test, WPAT: Word Place Association Test, K-TMT-E: Korean-Trail Making Test-Elderly, RCFT: Rey Complex Figure Test, SVLT: Seoul Verbal Learning Test. (continued to the next page)

https://doi.org/10.12779/dnd.2025.24.4.286

After adjustment for covariates, SCST subtests, domains, and the composite score showed stepwise decrements from CU to MCI to dementia (overall $p<0.001$). Key pairwise contrasts remained significant after Bonferroni correction (**Fig. 2**, **Supplementary Tables 3-5**).

AUC of the SCST composite score for the final clinical diagnosis

In covariate-adjusted ROC analyses (adjusted for age, sex, and education) in the pooled sample (n=777), the SCST composite score demonstrated high discriminative ability: AUC=0.903 (95% confidence interval [CI], 0.882–0.925) for CU vs. CI, 0.980 (95% CI, 0.970–0.991) for CU vs. dementia, 0.854 (95% CI, 0.823–0.886) for CU vs. MCI, 0.883 (95% CI, 0.855–0.911) for MCI vs. dementia, and 0.916 (95% CI, 0.894–0.937) for no dementia vs. dementia (**Fig. 3A**). Cohort-stratified AUCs were comparable: in the SCST–CERAD cohort the AUCs (95% CIs) were 0.960 (0.932–0.988), 0.994 (0.985–1.000), 0.926 (0.874–0.978), 0.904 (0.842–0.965), and 0.951 (0.919–0.983), respectively (**Fig. 3B**); in the SCST–SNSB cohort the AUCs (95% CIs) were 0.891 (0.865–0.918), 0.978 (0.966–0.991), 0.840 (0.802–0.878), 0.879 (0.847–0.911), and 0.907 (0.882–0.932), respectively (**Fig. 3C**). Detailed numerical estimates and other diagnostic indices are provided in **Table 2**.

B

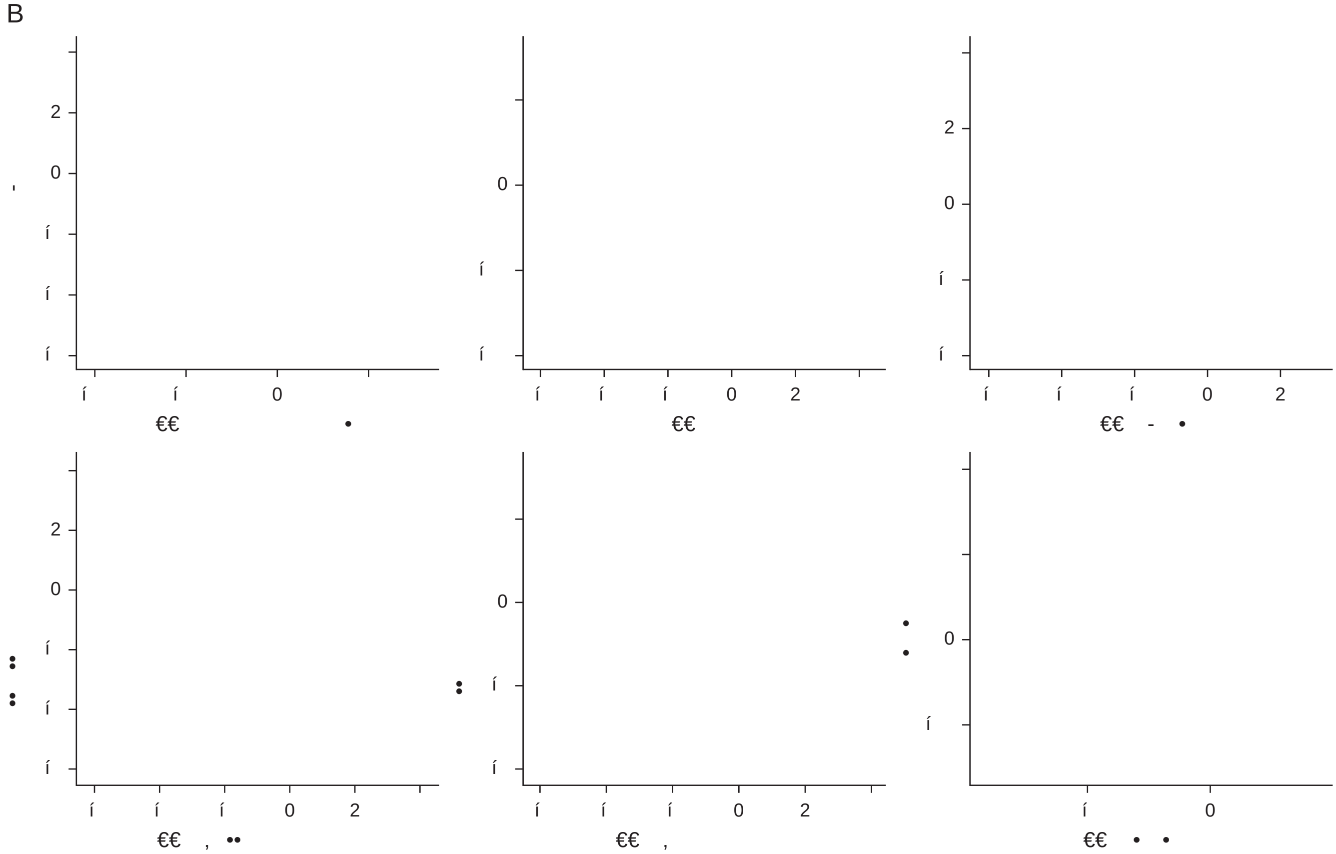

Fig. 1. (Continued) Correlation scatterplots between the SCST subtests (y-axis) and corresponding subtests of the CERAD-K and SNSB-II (x-axis). (A) SCST-CERAD cohort, (B) SCST-SNSB cohort. Scatterplots are based on standardized scores for visualization.
SCST: Seoul Cognitive Status Test, CERAD-K: Korean version of CERAD, SNSB-II: second edition of SNSB, CERAD: Consortium to Establish a Registry for Alzheimer's Disease, SNSB: Seoul Neuropsychological Screening Battery, DNT: Difficult Naming Test, K-BNT: Korean version of the Boston Naming Test, WPAT: Word Place Association Test, K-TMT-E: Korean-Trail Making Test-Elderly, RCFT: Rey Complex Figure Test, SVLT: Seoul Verbal Learning Test.