**Appendix**

**Wechsler Abbreviated Scale of Intelligence- 2nd Edition (WASI-II).** The WASI-II ([Wechsler, 2011](#_ENREF_43)), a revision of the WASI ([Wechsler, 1999](#_ENREF_42)), is a brief cognitive test designed to assess examinees aged six to 90 years. In addition to estimating general intellect, or a Full Scale IQ (FSIQ–4 and FSIQ–2 Subtest), it also provides composite scores estimating verbal skills (Verbal Comprehension; VCI) and visuospatial reasoning abilities (Perceptual Reasoning Index; PRI). It comprises four subtests: Vocabulary (verbal expression and word knowledge) and Similarities (verbal concept formation) subtests, which combine to yield the VCI, and Block Design (visuo-constructional skills) and Matrix Reasoning (abstract reasoning) subtests, which combine to yield the PRI. Instructions for all subtests were translated to Arabic, while modifying Vocabulary and Similarities items to the Saudi culture. Vocabulary and Similarities were piloted on 15 neurologically healthy controls, aged 16 -78. Items were then reordered based on difficulty, and the ones with poor response rate, or a level of difficulty that did not match the original English difficulty were replaced (i.e. English words with great sacrifice to level of difficulty after translation were replaced with new Arabic words with lower frequency through visual searches of synonyms in Arabic dictionaries).

**Jeddah Adaptation of the Boston Naming Test (JABNT).**The JABNT is a modification of the Boston Naming Test (BNT; [Kaplan, Goodglass, & Weintraub, 1983](#_ENREF_20), [2001](#_ENREF_21)), adapted by the first and the last authors to the Saudi culture. The BNT, one of the most widely used confrontation-naming task and known for its sensitivity to temporal lobe epilepsy ([Loring, 2010](#_ENREF_25); [Loring et al., 2008](#_ENREF_26); [Mayeux, Brandt, Rosen, & Benson, 1980](#_ENREF_30); [Schefft, Testa, Dulay, Privitera, & Yeh, 2003](#_ENREF_36)), consists of 60 line drawings of objects. The examinee attempts to name the drawings, and is given semantic and phonemic cues when applicable. It had been adapted to many languages, including Korean ([Kim & Na, 1999](#_ENREF_22)), Greek ([Patricacou, Psallida, Pring, & Dipper, 2007](#_ENREF_34)), French-Swiss ([Thuillard-Colombo & Assal, 1992](#_ENREF_40)), and Italian ([Riva, Nichelli, & Devoti, 2000](#_ENREF_35)) ([Fernández, 2013](#_ENREF_16)). For the JABNT, a total of 33 items were retained from the BNT, and 32 new items were added from [Snodgrass and Vanderwart (1980)](#_ENREF_38) to suit the Saudi culture. Items were initially ordered according to the authors’ estimation of difficulty for Saudi Arabian speakers (e.g. instead of being item # 17 in the BNT, “camel” was moved to item # 4 due to its commonality for Arabs; whistle was moved from 5 in the BNT to 26 in JABNT for proposedly lower familiarity). The 65 items were then piloted on 15 neurologically and healthy Saudi normal controls, aged 16-78. After that, items were reordered based on difficulty, demonstrated by response rate on the pilot study. Items with the lowest response rate and poorest response consistency were removed. Sixty items were eventually retained to form the JABNT.

**Bakker-Brandt Naming Test (BBNaT).** The BBNat ([Brandt, Bakker, & Maroof, 2010](#_ENREF_7)) consists of two subscales. The auditory naming task requires identification of the source of environmental sounds (i.e., animal calls, vehicles, instruments), and offers multiple-choice recognition for unidentified items. It provides a comparable visual naming task for auditory-visual comparison. The BBNat was shown to discriminate Alzheimer’s disease patients from normal participants and had acceptable internal consistency. The BBNaT items and instructions were translated to Arabic and back-translated to English, and culturally appropriate changes were made when necessary.

**Letter and animal based word fluency.** The wordlist generation test, commonly known as the verbal fluency task, measures spontaneous retrieval and efficient organization of words under restricted search conditions. It taps on several executive and other cognitive functions, including initiation, speed, long term vocabulary storage, ability to maintain mental set, shifting, monitoring, and inhibition ([Lezak, Howieson, Bigler, & Tranel, 2012](#_ENREF_24); [Mitrushina, Boone, Razani, & D'Elia, 2005](#_ENREF_32); [Strauss, Sherman, & Spreen, 2006](#_ENREF_39)). On the letter-based generation task (also known as the controlled word association test; COWA) the examinee is expected to name as many words that starts with a certain letter within a period of one minute. On the category-based wordlist generation (also known as semantic fluency), the examinee is asked to name as many words as possible that belong to a certain category (e.g. animals) within one minute. The Arabic word production tasks developed by Khalil (2010) were used for the Arabic battery in this study. These include both letter-based fluency (WRG) and category based (animal) fluency.

**Token Test.** The Token Test is a subtest of the Multilingual Aphasia Examination ([MAE; Benton, Hamsher, & Sivan, 1994](#_ENREF_6)) which assesses the presence, severity, and type of aphasia. It taps on verbal comprehension through the ability to follow 22 commands at two relative levels of difficulty, simple and complex, with excellent validity ([Strauss et al., 2006](#_ENREF_39)). The test was translated to Arabic and back-translated to English, with very minimal culturally required changes in wording.

**Baltimore Board.** The Baltimore Board ([previously known as the Hopkins Board; Brandt, Shpritz, Munro, Marsh, & Rosenblatt, 2005](#_ENREF_9)) was intended to be a relatively culture-free naming, learning, and content and location memory task. It was developed by the last author through a process of selecting nine picture stimuli common in all cultures (e.g. body parts, sun, tree), as deemed universally recognizable and namable by a panel of 10 clinical psychologists. The test proved particularly useful in the clinical assessment of temporal lobe epilepsy at Johns Hopkins. After naming them, the nine picture cards are placed in predetermined positions on a 35.56 x 35.56 cm board on which a 3 x 3 grid was drawn. The examinee is asked to watch the examiner place the pictures on the board, and remember their locations. The examinee then attempts to place them in their original positions. Incorrect placements are scored as errors and are immediately corrected. This procedure is repeated until two errorless consecutive trials are achieved. After 20-30 minutes, the examinee is asked to recall the names and the locations of the nine cards. Test items and instructions were translated to Arabic and back-translated to English, with culturally appropriate changes in wording.

**Logical Memory subtest of the Wechsler Memory Scale – Third Edition (WMS-III).**This subtest consists of a set of two stories used as memory stimuli. The passages are read to the examinee, and he or she is required to repeat them immediately, and again after a 25-30-minute delay. It also has a delayed recognition-yes/no-testing component, and has shown ability to identify left temporal lobe focus ([Baxendale et al., 1998](#_ENREF_4); [Lencz et al., 1992](#_ENREF_23); [Martin et al., 1999](#_ENREF_28)). A Saudi modification of the WMS-III ([Wechsler, 1997](#_ENREF_41)) was adapted and validated in Saudi Arabia primarily using clinical samples ([Escandell, 2002](#_ENREF_15); [Hassan, 2012](#_ENREF_19)). The Saudi WMS-III failed in localizing brain dysfunction (yet its auditory-visual discrepancy score was able to successfully differentiate left from right temporal lobe dysfunction). The two subgroups also differed significantly from one another in eight of the ten WMS-III Subtest scores ([Hassan, 2012](#_ENREF_19)) including Logical Memory. The subtest was therefore used in this study. An alternate form was created by the first and last authors, since the original form was in use in clinical evaluations already in place at KFSRHC at the time of this study.

In Escandell and Hassan’s version, the original theft story was maintained, with few cultural changes: the main character’s name “Anna Thompson” was changed to “Fatima Alzahrani,” both first and last names are very common in Saudi Arabia. Ms. Alzaharani is a vendor who sells spices on a known street in Riyadh, rather than a school lunch lady. The latter job nearly did not exist in the Saudi culture as occupations for Saudi women at the time were largely focused to teaching (more privileged) and vendor selling (for the less privileged). The first story in the alternate form also largely resembled a story from older versions of the WMS, in which a truck breaks down in the early morning hours. Main changes, implemented to provide a more updated feel, included that the man was carrying eggs and traditional Saudi bread in his jeep before it went off road, and that his phone ran out of battery before finding a backup charger and being able to call for help.

**Hopkins Verbal Learning Test-Revised (HVLT-R).** Among neuropsychological tests, wordlist memory tasks have demonstrated robust validity in temporal lobe epilepsy evaluations ([Davies, Bell, Bush, & Wyler, 1998](#_ENREF_12); [Loring et al., 2008](#_ENREF_26)). The HVLT-R ([Brandt & Benedict, 2001](#_ENREF_8)) was developed to allow brief repeated memory assessment of patients with dementia and those who cannot undergo lengthy assessment. It is a 12-item wordlist drawn equally from three semantic categories. The wordlist is read to the examinee three times and after each trial, the examinee is required to repeat as many words as possible. After 20-25 minutes, a delayed free recall trial is administered. This is followed by a 24-word-yes/no-recognition-trial consisting of 12 targets and 12 distracters (six of which are semantically related to the targets). The HVLT-R has established test retest reliability, as well as construct, concurrent and discriminant validity ([Shapiro, Benedict, Schretlen, & Brandt, 1999](#_ENREF_37)). It is available in six equivalent forms of similar psychometric properties. All forms and instructions were translated to Arabic, culturally appropriate changes were made under the supervision of the last author (also developer of the original test), and back-translated to English by the same back translators, with additional changes made when necessary. However, forms 2, 3, 4, and 6 were deemed culturally inappropriate to Saudi individuals due to semantic categories containing the following, respectively: alcoholic beverages (prohibited in the Islamic faith), musical instruments (not widely taught or encouraged at Saudi schools in the past 35 years), carpenter’s tool (labor-type occupations were largely taken upon non-Saudis in the past), and fish types (lack of rivers and only small percentage of Saudis live on the coasts, thus some fish types maybe unfamiliar to nomadic Saudis or those living in the mountainous South). Forms 1 and 5 were used in this study. In form 1, we used the “precious stones” category as in the English version, though “sapphire” (simply translated to “blue ruby” or the Arabic word “exhale”) was replaced by “agate,” and “opal” by “turquoise” as they were more common in the Arab culture, and allowed for less phonetic confusion. No content change was made to form 5.

**Brief Visuospatial Memory Test—Revised (BVMT-R).** The BVMT-R ([Benedict, 1997](#_ENREF_5)) is a figural learning and memory test that utilizes a multiple-trial list-learning paradigm. It assesses visuospatial immediate recall, learning curve, delayed recall, retention, and recognition. After studying six geometric figures for 10 seconds (displayed on an 8 x 11-inch sheet), the examinee is asked to reproduce as many accurate figures as possible, and in the same location they appeared in the matrix. This is repeated two more times using the same display. After 25 minutes, the examinee is to reproduce the designs. This is followed by a yes/no-recognition trial in which the examinee is shown 12 designs one at a time (six targets and six non-targets). On a number of clinical groups, validity was good, as was reliability ([especially inter-rater reliability; Benedict, 1997](#_ENREF_5)). The BVMT-R consists of six alternate forms with the same instructions; the first one was used in this study. Instructions were translated to Arabic by Psychological Assessment Resources, Inc (PAR) and the first author, and back-translated by the study’s two back translators.

**Color Trails Test (CTT).** The CTT ([D'Elia, Satz, Uchiyama, & White, 1996](#_ENREF_11)) is similar to the known Trail Making Test in that it assesses visual attention and scanning, processing speed, sequencing, mental flexibility, and motor function. However, the CTT does not use the English alphabet. The CTT was created in response the World Health Organization’s request to minimize language influences of the TMT and allow for cross-cultural application. On the first part (CTT-1), the examinee is asked to connect numbers from 1 to 25 in the correct order. On the second part (CTT-2), he or she is required to do the same, however, alternate between pink and yellow circles, disregarding the numbers in circles of the same color each time. Reliability and validity for the CTT with adults was reported as acceptable. For this study, Arabic numerals in English were replaced by Arabic numerals in Arabic language (the latter are interestingly of Indian origins). Instructions were translated to Arabic and back-translated to English, with minimal change, and reviewed and adjusted by one of the original test authors.

**Stroop Color-Word Test.** Several studies have used variation of the Stroop test with Arabic native speakers and bilinguals ([Al-Ghatani, Obonsawin, & Al-Moutaery, 2010](#_ENREF_1); [Alansari & Baroun, 2004](#_ENREF_2); [Bahri & Bendania, 1997](#_ENREF_3)). These studies, have generally demonstrated various performance differences between Arabic-English bilingual speakers, with modest information on validity with Arabs thus far. An Arabic, three-color modification of this test ([Al-Ghatani et al., 2010](#_ENREF_1)), based on the format by [Golden and Freshwater (2002)](#_ENREF_17), was used in this study.

**Grooved Pegboard Test (GPT).**The GPT ([Matthews & Klove, 1964](#_ENREF_29)) is a brief measure of manual dexterity and motor speed. Discrepancies from age-adjusted normative scores were hypothesized to lateralize cerebral dysfunction.The GPT consists of 25 pegs and a 10-centimeter-square metal board containing an array of randomly slotted holes. Using one hand at a time, the respondent is asked to lift and rotate each peg to fill in the holes row-by-row. The examiner records time of completion and number of drops. Test-retest reliability for the GPT has been mostly acceptable ([Dikmen, Heaton, Grant, & Temkin, 1999](#_ENREF_13)) and several studies supported its validity (e.g., [Haaland & Delaney, 1981](#_ENREF_18); [Mahurin & Inbody, 1989](#_ENREF_27)). Instructions for this test was translated to Arabic and back-translated according to the methodology described above.

**Edinburgh Handedness Inventory (EHI).** The EHI ([Oldfield, 1971](#_ENREF_33)) is a widely used instrument that assesses hand preference on 10 activities. For each activity, the respondent indicates his or her hand preference by putting a checkmark (or two when having a strong preference) in the appropriate column (right or left). When indifferent, the respondent is asked to put a checkmark in both columns. The EHI provides a laterality quotient that ranges from +100 (totally right handed) to -100 (totally left handed). After translation of instructions and items, the seven-item scale of the EHI was chosen for this study as it eliminates items with redundancy and error variance ([Dragovic, 2004](#_ENREF_14); [Milenkovic & Dragovic, 2013](#_ENREF_31)). Also, the seven-item scale has significantly better latent construct validity as indicated by factor regression analysis.

**Arabic Symptom Checklist-Revised (SCL-90-R).** Elbehairy (2004) first translated, into Arabic, and normed the SCL-90 in 1984; he then published the translated and normed Arabic Symptom Checklist-Revised in 2004. The checklist is a self-report questionnaire that measures symptoms of for depression, anxiety, phobia, paranoia, psychoticism, somatization, obsessive-compulsive tendencies, hostility, and interpersonal sensitivity. The questionnaire is in formal Arabic (Fus’ha) and Egyptian norms are available. Internal consistency and test retest reliability were established, and convergent validity was moderate.

**Quality of Life for Epilepsy Inventory-31** **(QoLiE-31).** QoLiE-31 ([Cramer et al., 1998](#_ENREF_10)) consists of seven scales that measure the following health aspects: Emotional wellbeing, social functioning, energy/fatigue, cognitive functioning, seizure worry, medication effects, and rated overall quality of life. It also provides an item (# 31) that assesses overall health condition, and a combined overall measure score that does not include the last item. Internal consistency and test retest reliability were adequate to high except for the medication side effect scale (*r* = .64). This inventory was translated to Arabic and back-translated according to the methodology described above.

References

Al-Ghatani, A. M., Obonsawin, M. C., & Al-Moutaery, K. R. (2010). Arabic version of Stroop test and its equivalency to the English version. *Pan Arab Journal of Neurosurgery, 14*(1), 112-115.

Alansari, B. M., & Baroun, K. (2004). Gender and cultural performance differences on the Stroop Color and Word test: A comparative study. *Social Behavior & Personality, 32*(3), 233-243. doi: 10.2224/sbp.2004.32.3.233

Bahri, T., & Bendania, A. (1997). Effect of language dominance on cognitive processes in a Stroop task. *Perceptual & Motor Skills, 84*(3), 979-988. doi: 10.2466/pms.1997.84.3.979

Baxendale, S., van Paesschen, W., Thompson, P., Connelly, A., Duncan, J. S., Harkness, W. F., & Shorvon, S. D. (1998). The relationship between quantitative MRI and neuropsychological functioning in temporal lobe epilepsy. *Epilepsia, 39*(2), 158-166.

Benedict, R. H. (1997). *Brief Visuospatial Memory Test—Revised*. Odessa, FL: Psychological Assessment Resources.

Benton, A. L., Hamsher, K. d., & Sivan, A. B. (1994). *Multilingual Aphasia Examination*. San Antonio, TX: Psychological Corporation.

Brandt, J., Bakker, A., & Maroof, D. (2010). Auditory confrontation naming in Alzheimer's disease. *The Clinical Neuropsychologist, 24*(8), 1326-1338.

Brandt, J., & Benedict, R. H. (2001). *Hopkins Verbal Learning Test–Revised*. Odessa, Fla: Psychological Assessment Resources.

Brandt, J., Shpritz, B., Munro, C. A., Marsh, L., & Rosenblatt, A. (2005). Differential impairment of spatial location memory in Huntington’s disease. *Journal of Neurology, Neurosurgery & Psychiatry, 76*(11), 1516-1519. doi: 10.1136/jnnp.2004.059253

Cramer, J. A., Perrine, K., Devinsky, O., Bryant-Comstock, L., Meador, K., & Hermann, B. (1998). Development and cross-cultural translations of a 31-item quality of life in epilepsy inventory. *Epilepsia, 39*(1), 81-88.

D'Elia, L. F., Satz, P., Uchiyama, C., L, & White, T. (1996). *Color Trails Test*. Odessa, FL: Psychological Assessment Resources.

Davies, K. G., Bell, B. D., Bush, A. J., & Wyler, A. R. (1998). Prediction of verbal memory loss in individuals after anterior temporal lobectomy. *Epilepsia, 39*(8), 820-828.

Dikmen, S. S., Heaton, R. K., Grant, I., & Temkin, N. R. (1999). Test-retest reliability and practice effects of expanded Halstead-Reitan Neuropsychological Test Battery. *Journal of the International Neuropsychological Society, 5*(4), 346-356.

Dragovic, M. (2004). Towards an improved measure of the Edinburgh Handedness Inventory: A one-factor congeneric measurement model using confirmatory factor analysis. *Laterality: Asymmetries of Body, Brain and Cognition, 9*(4), 411-419.

Escandell, V. A. (2002). Cross-cultural neuropsychology in Saudi Arabia. In F. R. Ferraro (Ed.), *Minority and cross-cultural aspects of neuropsychological assessment.* (pp. 299-325). Lisse, Netherlands: Swets & Zeitlinger Publishers

Fernández, A. L. (2013). Development of a confrontation naming test for Spanish-speakers: the Cordoba Naming Test. *The Clinical Neuropsychologist, 27*(7), 1179-1198. doi: 10.1080/13854046.2013.822931

Golden, C. J., & Freshwater, S. M. (2002). *Stroop Color and Word Test: Revised examiner’s manual.* Wood Dale, IL: Stoelting Co.

Haaland, K. Y., & Delaney, H. D. (1981). Motor deficits after left or right hemisphere damage due to stroke or tumor. *Neuropsychologia, 19*(1), 17-27.

Hassan, A. M. (2012). Neuropsychological evaluation in epilepsy surgery – A cross-cultural perspective. In D. Stevanovic (Ed.), Epilepsy: Histological, electroencephalographic and psychological aspects (pp. 175-198): InTech. Retrieved from <http://www.intechopen.com/books/>.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia: Lea & Febiger.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test*. Philadelphia: Lea & Febiger.

Kim, H., & Na, D. L. (1999). Normative data on the Korean version of the Boston Naming Test. *Journal of clinical and experimental neuropsychology, 21*(1), 127-133. doi: 10.1076/jcen.21.1.127.942

Lencz, T., McCarthy, G., Bronen, R. A., Scott, T. M., Inserni, J. A., Sass, K. J., . . . Spencer, D. D. (1992). Quantitative magnetic resonance imaging in temporal lobe epilepsy: Relationship to neuropathology and neuropsychological function. *Annals of Neurology, 31*(6), 629-637. doi: 10.1002/ana.410310610

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment (5th ed.)*. New York, NY, US: Oxford University Press.

Loring, D. W. (2010). History of neuropsychology through epilepsy eyes. *Archives of Clinical Neuropsychology, 25*(4), 259-273. doi: 10.1093/arclin/acq024

Loring, D. W., Strauss, E., Hermann, B. P., Barr, W. B., Perrine, K., Trenerry, M. R., . . . Bowden, S. C. (2008). Differential neuropsychological test sensitivity to left temporal lobe epilepsy. *Journal of the International Neuropsychological Society, 14*(3), 394-400. doi: 10.1017/S1355617708080582

Mahurin, R. K., & Inbody, S. B. (1989). Psychomotor assessment of the older patient. *Clinics in Geriatric Medicine, 5*(3), 499-518.

Martin, R. C., Hugg, J. W., Roth, D. L., Bilir, E., Gilliam, F. G., Faught, E., & Kuzniecky, R. I. (1999). MRI extrahippocampal volumes and visual memory: correlations independent of MRI hippocampal volumes in temporal lobe epilepsy patients. *Journal of the International Neuropsychological Society, 5*(6), 540-548.

Matthews, C. G., & Klove, K. (1964). *Instruction manual for the Adult Neuropsychology Test Battery*. Madison, WI: University of Wisconsin Medical School.

Mayeux, R., Brandt, J., Rosen, J., & Benson, D. F. (1980). Interictal memory and language impairment in temporal lobe epilepsy. *Neurology, 30*(2), 120-125.

Milenkovic, S., & Dragovic, M. (2013). Modification of the Edinburgh Handedness Inventory: A replication study. *Laterality: Asymmetries of Body, Brain and Cognition, 18*(3), 340-348. doi: doi:10.1080/1357650X.2012.683196

Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment (2nd ed.)*. New York, NY, US: Oxford University Press.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia, 9*, 97-113.

Patricacou, A., Psallida, E., Pring, T., & Dipper, L. (2007). The Boston Naming Test in Greek: Normative data and the effects of age and education on naming. *Aphasiology, 21*(12), 1157-1170.

Riva, D., Nichelli, F., & Devoti, M. (2000). Developmental aspects of verbal fluency and confrontation naming in children. *Brain and Language, 71*(2), 267-284. doi: 10.1006/brln.1999.2166

Schefft, B. K., Testa, S. M., Dulay, M. F., Privitera, M. D., & Yeh, H. S. (2003). Preoperative assessment of confrontation naming ability and interictal paraphasia production in unilateral temporal lobe epilepsy. *Epilepsy & Behavior, 4*(2), 161-168.

Shapiro, A. M., Benedict, R. H., Schretlen, D., & Brandt, J. (1999). Construct and concurrent validity of the Hopkins Verbal Learning Test-Revised. *The Clinical Neuropsychologist, 13*, 348–358.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology. Human Learning & Memory, 6*(2), 174-215.

Strauss, E., Sherman, E., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary (3rd. ed)*. New York, NY, US: Oxford University Press.

Thuillard-Colombo, F., & Assal, G. (1992). Adaptation française du test de dénomination de Boston versions abrégées. *Revue Européenne de Psychology Appliquée, 42*, 67–71.

Wechsler, D. (1997). *WMS-III Administration and Scoring Manual*. San Antonio, Texas: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Intelligence Scale*. San Antonio, Texas: The Psychological Corporation.

Wechsler, D. (2011). *Wechsler Abbreviated Intelligence Scale, Second Edtion*. San Antonio, Texas: The Psychological Corporation.