Reliability of a United States Version of the Nottingham Sensory Assessment

May 2015

This research, submitted by Amy Kaestner and Rebecca Miller, has been approved and accepted in partial fulfillment of the requirements for the degrees of Master of Science in Occupational Therapy from the University of Puget Sound

Committee Chairperson: Sue Doyle, PhD, OTR/L

Director, Occupational Therapy Program: Yvonne Swinth, PhD, OTR/L, FAOTA

Dean of Graduate Studies: Sunil Kukreja, PhD

KEY WORDS: sensory assessment, inter-rater reliability, United States Nottingham Sensory Assessment (US-NSA)
Abstract

Many stroke survivors experience somatosensory deficits and there is currently no “gold standard” reliable standardized assessment commonly used by clinicians in the United States. In the present study, the authors modified the Nottingham Sensory Assessment (NSA) into a U.S. version to provide therapists with a standardized multimodal sensory assessment for use with clients post-stroke. Six licensed rehabilitation practitioners and one occupational therapy student administered the United States-NSA (US-NSA) on 17 older adults with chronic stroke (i.e., at least six months post-stroke) to evaluate its inter-rater reliability. The authors used an intraclass correlation coefficient (ICC) to analyze the inter-rater reliability of the data. The findings indicate strong agreement for all test items, except for the Sharp-Dull Discrimination item when tested on the lower extremity. The US-NSA is a promising sensory impairment assessment available for use by U.S. occupational therapists and other rehabilitation professionals to use in intervention planning, measuring outcomes, and quality of care.
Reliability of a United States Version of the Nottingham Sensory Assessment

Somatic sensation enables humans to engage with the world in a safe and meaningful way. It supports a spectrum of skills that underlie participation in numerous areas of occupation, from detecting dangerous levels of heat to knowing where one’s body is in space. Nearly 800,000 Americans suffer a new or recurrent stroke every year (American Heart Association, 2013), and more than half of all stroke survivors experience sensory dysfunction (Carey, 1995). These sensory deficits can dramatically impact a person’s ability to function and live safely and independently. Studies suggest that sensory impairment is both related to motor impairment (Tyson, Hanley, Chillala, Selley, & Tallis, 2008) and predictive of motor recovery after stroke (Feys et al., 2000). There is potential that addressing sensation in the intervention process may improve mobility, motor abilities, and functional independence in stroke survivors.

Clinicians working with stroke survivors report that they use sensory assessments on a regular basis and find them important for clinical practice (Doyle, Bennett, Gustafsson, 2013b; Winward, Halligan, & Wade, 1999). Information gained from sensory assessments is important for the entire therapeutic process: diagnosis, prognosis, client and caregiver education, and intervention planning (Winward et al., 1999). It follows that being able to determine the extent and severity of sensory deficits in stroke survivors allows for more informed treatment and may potentially lead to better client outcomes (Winward et al., 1999). The American Occupational Therapy Association (AOTA) (2010) strongly recommends that occupational therapists use appropriate, evidence-based measures during the evaluation and intervention processes as much as possible in order to guide effective clinical decision-making. There are a limited number of standardized sensory measures available, however, and most possess low clinical utility and have not been found to be reliable or valid (Connell & Tyson, 2012).

The Nottingham Sensory Assessment (NSA) (Lincoln et al., 1991), originally developed in Europe, is a standardized multimodal sensory assessment that has been revised for use in many countries. Subsequent revisions of the NSA are currently available across Europe (Lincoln, Jackson, & Adams, 1998; Stolk-Hornsveld, Crow, Hendriks, Van Der Baan, & Harmeling-Van Der Wel, 2006) and in Brazil.
The Erasmus MC modifications to the Nottingham Sensory Assessment (EmNSA) has demonstrated strong clinical utility and high psychometric values compared to other standardized sensory assessments currently available to occupational therapists (Connell & Tyson, 2012). The EmNSA assesses a wide range of sensory modalities, tactile detection and discrimination as well as proprioception, on both the upper and lower limbs (Stolk-Hornsveld et al., 2006). It offers ease of use in a clinical setting because it is fast to administer, requiring between 10 and 15 minutes, and involves inexpensive and readily available testing materials (Stolk-Hornsveld et al., 2006). Stolk-Hornsveld et al. (2006) has found that the inter-rater and intra-rater reliability for the EmNSA ranges from good to excellent across different components of the assessment. Although this assessment was found to have one of the best balances between reliability and clinical utility (Connell & Tyson, 2012), the EmNSA is rarely used by U.S. occupational therapists (Doyle, Bennett, & Gustafsson, 2013b). U.S. clinicians choose to make use of less reliable and less comprehensive assessments (Doyle et al., 2013b), which provides therapists with unreliable information for intervention planning and clinical decision-making. One potential reason for the limited use of the NSA in the United States could be that there are unique differences in language and culture between Europe and the United States. These dissimilarities are significant enough to potentially cause U.S. occupational therapists and patients to misunderstand some of the test items, which may reduce the reliability of the NSA when used in the United States.

**Background**

**Characteristics of stroke.** Stroke affects hundreds of thousands of Americans each year making it the fourth leading cause of death in America (American Heart Association, 2013). Stroke occurs as a result of a bleed or decreased blood supply to the brain leading to acute brain damage. The most common damage seen in persons post-stroke is motor deficiencies on the contralesional side of the body (Rand, Gottlieb, & Weiss, 2001). Although the severity and nature of the damage varies from person to person, restoring motor deficiencies is a crucial aspect of stroke rehabilitation because motor abilities underlie functional activities (Tyson et al., 2008).

Studies have found that 53-89% of persons post-stroke experience sensory deficits (Acerra, 2007;
Rand et al., 2001; Tyson et al., 2008), and the extent and severity of sensory damage also varies from client to client. Sensory impairments diminish the amount of feedback the body can receive from the environment. Pain receptors alert one to the presence of dangerous stimuli. Proprioceptive input informs one about where the body is in space in order for one to plan and execute voluntary movements. Several studies have found that sensory dysfunction after stroke contributes to poorer functional outcomes (Rand et al., 2001; Sullivan & Hedman, 2008; Tyson et al., 2008). Further, some studies have found evidence that function improves following intervention that includes sensory retraining (Byl et al., 2003; Celnik, Hummel, Harris-Love, Wolk, & Cohen, 2007; Floel et al., 2004; Smania, Montagnana, Faccioli, Fiaschi, & Aglioti, 2003). As such, addressing sensory function should be another critical aspect of stroke rehabilitation.

**Current practice patterns.** International clinical practice guidelines have recommended the use of standardized assessment tools for evaluating sensory deficits after stroke. In particular, both the British and Australian practice guidelines suggest using the NSA in stroke rehabilitation (Doyle, 2015). The U.S. clinical practice guidelines, however, have not made specific recommendations regarding the use of standardized assessment tools (Doyle, 2015). As such, there is currently no accepted gold standard in the United States for evaluating sensory deficits after stroke. The assessments that are available to occupational therapists do not have standardized methodologies, possess low clinical utility, or have weak psychometric properties (Connell & Tyson, 2012; Doyle, 2015; Doyle et al., 2013b; Lin, Hsueh, Sheu, & Hsieh, 2004; Sullivan & Hedman, 2008). The “standard neurological evaluation” for examining sensory deficits consists of single-sense examinations that test discrete modalities and focus primarily on the modalities of light touch, pain, and temperature (Doyle, 2015; Doyle, Bennett & Gustafsson, 2013b). Other modalities that are sometimes tested include proprioception, stereognosis, vibration, and two-point discrimination (Theis, 2014). Although the “standard neurological evaluation” is available and is called “standard,” it is not standardized nor has it been proven reliable (Dannenbaum, Michaeelsen, Desrosiers, & Levin, 2002). Dannenbaum et al. (2002) asserted that a lack of standardization is present among the sensory tests typically used to examine patients post-stroke. Many commonly used assessments leave the
selection of materials, locations of the body, number of trials, and modalities tested to the discretion of the examiner (Theis, 2014), which leads to inconsistent testing procedures and decreased reliability of the results.

A few standardized assessments are available, but many have low clinical utility, poor psychometrics, or both, which make them undesirable in clinical practice (Connell & Tyson, 2012; Sullivan & Hedman, 2008). One such assessment is the Rivermead Assessment of Somatosensory Perception. While this assessment has been widely reported as being used in research studies and demonstrates good inter-rater and intra-rater reliability (Winward, Halligan, & Wade, 2002), the materials necessary to administer this test are currently not available for purchase (Doyle, 2015). Another standardized assessment is the sensory subscale of the Fugl-Meyer Assessment, which is used to evaluate light touch and proprioceptive sensations (Lin et al., 2004). This subscale, however, has demonstrated a wide range of inter-rater reliability from poor to excellent using weighted kappa ($K_w$) coefficients ($K_w = .30$ to $K_w = .90$), for different test items making it unreliable for accurately evaluating sensory deficits (Lin et al., 2004). The authors of one systematic review concluded that the Fugl-Meyer Assessment had high construct validity and high clinical utility (Connell & Tyson, 2012). Many other authors, however, dispute its reliability and validity, and these inconsistent ratings negate other positive features of the Fugl-Meyer and do not support its clinical use as a sensory evaluation tool (Lin et al., 2004; Sullivan & Hedman, 2008).

It appears that the current manner in which U.S. therapists assess sensation does not comply with the recommendations from the OT practice guidelines in that therapists are not using standardized, reliable, and valid measures (Doyle, 2015). Nevertheless, occupational therapists perform and value information gained from sensory measures to inform intervention planning with patients post-stroke (Doyle et al., 2013b; Winward et al., 1999). The results of a survey, with a final sample of 145 occupational therapists from the American Occupational Therapy Association’s (AOTA) physical disability special interest group, found that fewer than 1% of occupational therapists reported using standardized multimodal assessments in practice with the majority reporting using components of the
“standard neurological evaluation” (Doyle et al., 2013b). This study received a relatively low response rate of 37% and collected data from a randomly selected sample of AOTA members (Doyle et al., 2013b). Consequently, the results may not be representative of all occupational therapists. Still, the survey revealed such a strong trend that it is safe to assert that U.S. therapists treating patients with sensory impairments generally do not use standardized sensory assessments. When examining the clinical decision-making of participants in this survey, these authors concluded that the majority of U.S. therapists use assessments that test single sensory modalities (Doyle, Bennett, & Gustafsson, 2013a). Limiting sensory evaluations to single sensory modalities may cause clinicians to miss other manifestations of sensory dysfunction. Kim and Choi-Kwon (1996) even found that the majority of patients diagnosed with no sensory deficits, as determined by the “standard neurological evaluation,” actually demonstrated sensory impairments when evaluated with a standardized multimodal sensory assessment. In addition, sensory impairments are not easily observable and therapists often rely on patient report to guide which sensory evaluations should be administered (Doyle et al., 2013a). Less than one fifth of stroke survivors report experiencing sensory deficits; yet when tested, more than half actually demonstrate such impairments (Acerra, 2007). It follows that patient report is not a reliable manner of determining the nature of sensory deficits or whether an evaluation of sensory impairment is warranted. Since therapists incorporate information gathered during sensory assessments into intervention planning (Winward et al., 1999), incomplete test data can undermine clinical decision-making regarding intervention planning, and may not adequately address underlying impairments. It is clear that therapists could benefit from a more comprehensive, standardized way of assessing sensory dysfunction.

In a time focused on outcomes and demonstrating the impact of therapy, occupational therapists are strongly encouraged to use reliable and valid standardized measurement tools. The Occupational Therapy Code of Ethics and Standards encourages occupational therapists to use evidence-based evaluation techniques whenever possible to inform clinical decision-making (AOTA, 2010). Non-standardized tests can impact evaluation accuracy, the ability to demonstrate client gains, and the quality
of therapy services. Standardized assessments enable clinicians to consistently measure and document progress and outcomes in order to deliver skilled services.

The current practice of using non-standardized single modality sensory measures not only affects client treatment, but it is also not in line with the AOTA Centennial Vision. According to the Centennial Vision, AOTA aspires to be science-driven as well as evidence-based by 2017 (AOTA, 2007). This goal will not be possible without high quality assessment tools that can be used in both clinical and research settings. Using the same standardized measurement tools across settings and research will increase clinicians’ ability to produce valuable research in the field (Rao, 2012) and enhance researchers’ ability to obtain and compare valuable results (Doyle, Bennett, Fasoli, & McKenna, 2010). Occupational therapy in the United States could benefit from a standardized multimodal sensory measure to use with patients post-stroke. Thus, it is important for high quality standardized multimodal measurement tools to be made available in the United States.

**Nottingham Sensory Assessment.** The NSA is an established standardized multimodal assessment used with patients post-stroke in other countries. The NSA was first developed in the U.K. in 1991 as a comprehensive sensory evaluation tool, and includes the following test items: Light Touch, Temperature, Pinprick, Pressure, Tactile Localization, Bilateral Simultaneous Touch, Stereognosis, Proprioception, and Two-Point Discrimination. Unfortunately, using Cohen’s kappa (K) coefficients, the assessment was found to have a wide range of inter-rater reliability (K = .01 to K = .89) (Lincoln et al., 1991). Several years later, Lincoln et al. (1998) revised the NSA into a shorter assessment. The researchers analyzed Light Touch, Pressure, Pinprick, Temperature, Kinesthetic Sense, and Two-Point Discrimination, and identified possible areas of overlap in the assessment. The researchers did not have enough data to analyze or make changes to the Stereognosis and Two-Point Discrimination sections. They then evaluated the reliability of the revised NSA (rNSA) and found that inter-rater reliability was acceptable (K = .04 to K = .77), and that use of the shorter assessment did not result in a loss of information (Lincoln et al., 1998). A decade later, Connell (2007) shortened the NSA through a Rasch analysis and found this shorter version to have high reliability and validity.
Gaubert and Mockett (2000) found the Nottingham method of stereognosis to have mostly good inter-rater reliability (K = .40 to K = .86). Stereognosis closely mirrors functional tasks that require the manipulation of objects in the hand as well as fine motor control without visual information (e.g., dressing, finding objects in purse, styling hair, bathing). Connell (2007) found that stereognosis was among the most significantly impaired sensations, occurring in 31–89% of hospital patients post-stroke that were recruited for her study (Connell, 2007). In addition, a later qualitative study found that stroke survivors reported considerable difficulty when performing tasks that required stereognosis ability (Doyle, Bennett, & Dudgeon, 2014). These findings suggest that developing a valid and reliable measure of stereognosis ability is important.

Stolk-Hornsveld et al. (2006) later modified the rNSA to create the EmNSA: a standardized sensory assessment that has predominantly good to excellent inter- and intra-rater reliability and only takes 10 to 15 minutes to administer. Recently the EmNSA was recommended for use by researchers who conducted a systematic review of sensory measures because of its combination of strong psychometric properties and clinical utility (Connell & Tyson, 2012). Temperature was excluded in this version, and Sharp-Blunt Discrimination was added (Stolk-Hornsveld et al., 2006). Using weighted kappa coefficients, the EmNSA has a range of mostly good to excellent inter-rater reliability (K_W = .46 to K_W = 1.00) and intra-rater reliability (K_W = .58 to K_W = 1.00) for subtests of Tactile Sensation (Light Touch, Pressure, Pinprick), Sharp-Blunt Discrimination, and Proprioception (Stolk-Hornsveld et al., 2006). The exception is the subtest of Two-Point Discrimination, which has poor to good agreement for intra-rater reliability (K_W = .11 to K_W = .63) and inter-rater reliability (K_W = .10 to K_W = .66) (Stolk-Hornsveld et al., 2006). The EmNSA measures multiple sensory modalities over all body sections in as little as 10 minutes without high cost equipment (Stolk-Hornsveld et al., 2006). It also offers ease of use with a wide range of patient conditions because it allows patients to respond with movements or cards if they are unable to verbalize answers (Connell, 2007).

**Proposed modifications to the NSA.** The NSA is rarely used in the United States (Doyle et al. 2013b), perhaps due to cultural and language differences as well as the considerable length of the original
RELIABILITY OF THE US-NSA

assessment. Modifying the text from British English to American English and combining the best versions of each test item would give U.S. therapists the opportunity to use a standardized multimodal sensory assessment to support the treatment of patients post-stroke and advance the literature about sensory loss. The revised NSA was successfully translated into Portuguese in 2010 and maintained good clinical utility and strong reliability (Lima et al., 2010). This suggests excellent potential for the NSA to be successfully modified for use in the United States.

The purpose of the present study is to develop a U.S. version of the NSA and evaluate its inter-rater reliability, as administered by licensed rehabilitation practitioners among adults with chronic stroke, in order to make a reliable, multimodal sensory impairment assessment available to U.S. occupational therapists.

Method

Research Design

This is a prospective study examining the inter-rater reliability of the United States-NSA (US-NSA). The design of this study was twofold. First, the test items included in previous versions of the NSA as well as the associated procedural instructions were modified from British English to American English to make the assessment more accessible and understandable to U.S. clinicians and clients. Secondly, experienced U.S. occupational therapists were recruited to evaluate persons with chronic stroke in order to examine the inter-rater reliability of the US-NSA. The goal of this study was to develop a clinical measurement tool that can be used by clinicians working with the stroke population in the United States. As such, conducting this study with practicing rehabilitation therapists increases the external validity because of the real-world replication and application. Testing persons with chronic stroke for this study was more appropriate than testing persons with acute stroke because the former have more stable sensory profiles, while the latter are more likely to experience spontaneous recovery and thus have less stable sensory profiles. The University of Puget Sound’s Institutional Review Board approved the procedures for this study.
Participants

The authors used a convenience sampling method to recruit patient and assessor participants. One population of interest for this study was persons with chronic stroke. Persons post-stroke were recruited via phone call. Phone numbers were obtained through the University of Puget Sound (UPS) onsite clinic record of current and past clients who had consented to be contacted for research studies. A consent form was given to those who were interested, eligible, and willing to participate in the study at the time of the first session. To be included as a patient participant in this study, persons had to be over the age of 18 with chronic (at least six months post-stroke) ischemic or hemorrhagic stroke.

The second population of interest for this study was licensed rehabilitation therapists. Clinical instructors at UPS were recruited for this study via email. These clinical instructors also recruited colleagues outside of the UPS community. A consent form was given to those who were interested, eligible, and willing to participate in the study at the time of the US-NSA training. To be included as an assessor participant, persons had to be licensed rehabilitation therapists with at least one year of experience in an adult physical dysfunction setting and have completed the training module on the US-NSA. The authors conducted a two-hour training workshop in which assessor participants were given training on the US-NSA administration procedures. A video demonstration of each US-NSA subtest was shown. A standardized set of instructions for the US-NSA was provided and assessors practiced the US-NSA in full with at least one attendee as part of the training. Assessors were given feedback during the practice and asked the authors questions to clarify the testing procedures. Successful completion of the entire training module was required to be eligible to participate as an assessor in the study. Assessor participants were excluded if they did not meet the above criteria or did not complete the full training module. Six therapists participated in this study to test inter-rater reliability.

Modifying the Assessment

The US-NSA was most closely based on the EmNSA scale because of its high reliability and clinical utility. Changes to the tool’s content and language were made through regular discussions among
the authors, an Australian/American occupational therapist and stroke researcher as well as two U.S.
occupational therapy student researchers.

The authors incorporated the recommendations made in the EmNSA study into the US-NSA. These recommendations were modified to account for language differences and material accessibility for the new assessment. Stolk-Hornsveld et al. (2006) retained three tests in the final version of the EmNSA, Tactile Sensation (Light Touch, Pressure, and Pinprick), Sharp-Blunt Discrimination, and Proprioception, based on their findings of predominantly good to excellent intra- and inter-rater reliability for these tests. Stolk-Hornsveld et al. (2006) found Two-Point Discrimination to have poor reliability and did not retain this test on the final score sheet; as such, it was also excluded from the US-NSA. Although the Stereognosis test was not included in the EmNSA study, it was added to the US-NSA because of the functional component of stereognosis sensation and its strong validity and reliability found in other previous editions of the NSA (Connell, 2007; Gaubert & Mockett, 2000; Lima et al., 2010).

The procedural instructions were modified to account for dialectical and cultural differences. These linguistic changes were agreed upon through regular discussions among the authors. A pilot review of the administration manual was conducted to gain feedback in order to ensure that the test and procedural instructions were easy to use and understandable to American English speakers. This pilot review was conducted with the UPS occupational therapy graduate students and expert clinicians. Nine second-year graduate students were asked to attempt to administer the US-NSA using only the administration manual and score sheet. They provided the authors with feedback regarding unclear directions and sections that were difficult to understand. After integrating this feedback, the authors presented the US-NSA to 30 first-year graduate students and two expert clinicians. The authors briefly demonstrated the testing procedures and students were given time to practice administering the measure independently. At the end of this session, the students and clinicians provided further feedback regarding the measure’s ease of use, method, and clarity of instructions. During the US-NSA training, the assessors felt that the addition of a script would be beneficial to further increase standardization and ease of use.
The authors integrated this feedback and developed a script that was used during testing, which can be found in Appendix A.

The materials needed for each test were examined to find counterparts easily understood and readily available in the United States. A prototypical example of a U.S. national brand option has been identified for each object and is provided in the US-NSA Administration Manual (Appendix B). These guidelines will allow clinicians to easily replicate the materials thereby maintaining standardization. The scoring scale is consistent with previous versions, although a new score sheet was created to match the language and assessment changes (Appendix A).

**US-NSA Instrumentation**

**Tactile Sensation.** Tactile Sensation was modified from the EmNSA version (Stolk-Hornsveld et al., 2006). It consists of three subtests to assess the patient's tactile detection. These three subtests are a) Light Touch, assessed using a cotton ball, b) Pressure, assessed using the index finger, and c) Pinprick, assessed using a Neurotip™. A list of defined points of contact can be found in Appendix B. For each component of the test, the assessor will stimulate each defined point of contact one time, in random order, and the patient will report whenever he or she can feel the test stimulus. The test administrator will record a score of 0 (Absent), 1 (Impaired) or 2 (Normal) for each body section (e.g., hand, forearm, thigh, foot) based on specific scoring criteria. If a score of 2 is achieved for all of a limb on the Light Touch subtest, a score of 2 is automatically assigned for all Pressure and Pinprick items for that limb (Stolk-Hornsveld et al., 2006). The scores are summed into four composite scores (right upper extremity, right lower extremity, left upper extremity, and left lower extremity) and each subtest (Light Touch, Pressure, Pinprick) is summed separately. Possible sum scores range from 0–8. Refer to the US-NSA score sheet in Appendix A.

**Sharp-Dull Discrimination.** The US-NSA’s Sharp-Dull Discrimination test was modified from the EmNSA version (Stolk-Hornsveld et al., 2006). Based on feedback from two expert reviewers, Sharp-Blunt Discrimination has been changed to Sharp-Dull Discrimination to match language commonly used by U.S. therapists to describe this sensory modality. Sharp-Dull Discrimination evaluates discrimination
between sharp and dull sensations with the use of a Neurotip™ and the tip of the index finger. This test is not administered, however, if the patient does not receive full scores for Tactile Sensation (Light Touch, Pressure, Pinprick). The stimulus is applied in a random order, at each defined point of contact (Appendix B) six times, three with the Neurotip™ and three with the index finger. The patient will describe or indicate the test sensation (e.g., sharp, dull) after each application of a stimulus. Scoring procedures for each body section and composite scores are consistent with those for Tactile Sensation.

**Proprioception.** The Proprioception test was modified from the EmNSA version (Stolk-Hornsveld et al., 2006). Increased detail to clarify hand placement was added to the instructions after feedback during US-NSA pilot testing. Changes to the sentence order and word choice were made to more closely imitate American English sentence structure. Proprioception requires the examiner to move the patient’s limbs through passive motions with defined starting points and hand placements (Appendix B). Movement at each joint is repeated three times. The patient must answer specific questions to demonstrate understanding of the direction of the movement that is taking place. If the patient is unable to describe the direction of the movement, he or she is asked instead to identify if the movement is occurring. The assessor then scores 0 (Absent), 1 (Impaired) or 2 (Normal) for each joint, and these scores are summed for left and right upper and lower extremities.

**Stereognosis.** The US-NSA also includes a Stereognosis test due to this test’s functional applications in daily life, which was modified from the rNSA version (Gaubert and Mockett, 2000). Connell (2007) found certain items in the rNSA Stereognosis test to be redundant based on her Rasch analysis. Her findings suggest that excluding these items from the testing procedures will increase the efficiency and efficacy of the Stereognosis test (Connell, 2007). She suggested that the “biro,” or pen, item be removed along with the plastic cup, scissors, and two of the three coins (Connell, 2007). One item used in the rNSA Stereognosis test was adjusted to account for differences in language (“flannel” became washcloth). The single U.K. coin retained from Connell’s analysis was replaced with a U.S. coin of similar ubiquity and recognition. Subsequently, the US-NSA Stereognosis test retains the following objects: quarter, pencil, comb, sponge, washcloth, and a glass.
The Stereognosis test requires the assessor to place everyday objects (e.g., comb, pencil, glass) one at a time into the patient’s hand for up to 30 seconds, and the patient must identify or describe each object with eyes occluded. For each object, the patient receives a score of 0 (Absent), 1 (Impaired) or 2 (Normal). Scores will be summed into a composite score ranging from 0-12. If the client is unable to move his/her hand to manipulate the object, the assessor may move the object around the patient’s fingers and palm. Refer to Appendix B for more information.

Two-Point Discrimination. The EmNSA method (Stolk-Hornsveld et al., 2006) of Two-Point Discrimination was excluded from the US-NSA. Stolk-Hornsveld et al. (2006) did not find that Two-Point Discrimination had high enough reliability to be viable in practice. Further, van Nes et al. (2008) found low sensitivity for this test in affected patients. Prescott, Garraway, and Akhtar (1982) found that assessing two-point discrimination had little prognostic power for return to function after stroke, which suggests that excluding this test will not result in a significant loss of information. Exclusion of this test will also increase clinical utility by reducing testing time and required testing materials.

Study Design

Each patient was rated two separate times by different assessors. Sessions were between two and eight days apart. Assessors were blind to patient results from previous evaluations prior to testing. Initially a testing schedule was developed to create different assessor combinations when evaluating patients to increase overlap and ensure that the same two assessors did not test the same group of patients. Due to complications in patient availability and changes in assessors’ schedules (e.g., illness, work), the authors were unable to follow this schedule fully. Two patients were unable to attend their original appointments, and could not make up sessions during a time that a therapist was available. As a result, one of the authors (AK), an occupational therapy student, performed the assessment with these two patients. This student was trained in US-NSA testing procedures and followed testing protocol. Details about the final testing distribution can be found in Table 1.

The US-NSA can be used on one side of the body as a screening tool or on both sides of the body as a more comprehensive assessment of sensory deficits. For the present study, each patient was tested on
only the more affected side of the body. The authors hypothesized that testing the more affected side would yield greater variability in sensory deficits and provide assessors with more diverse scoring opportunities. They believed this would better measure the assessor’s ability to score distinct sensory deficits in a reliable manner.

During evaluation sessions, assessors were located in a private room with the US-NSA assessment toolkit, a plinth, and two chairs. One at a time, patients entered the testing room and were positioned in supine, either independently or with assistance, on the plinth for the assessment. The assessor administered the US-NSA according to the standardized instructions and recorded the patient’s sensation using the standardized US-NSA score sheet. The score sheet included the participant ID number and the side of the body assessed, but names and other personally identifying information were not included. Data was recorded on the US-NSA standardized score sheet in order to accurately compare first and second test scores. The authors collected the score sheets from the assessor after each individual evaluation session.

Data Analysis

The data were entered into Statistical Package for Social Sciences (SPSS) at the completion of the trials. The authors then used intraclass correlation coefficients (ICCs) to analyze the inter-rater reliability of the data. ICCs (r) are commonly used to measure inter-rater reliability and some research suggests that it is appropriate to use with ordinal data with structured intervals between scores (Tinsley & Weiss, 1975).

Assessor participants were a sample of a larger population of raters and these results attempt to generalize to a single rater’s ability to use the test reliably in practice, indicating the need for an ICC (1, 1): a one way random single measures ICC (Shrout & Fleiss, 1979). The ICCs were interpreted as follows: ICC < .40 indicates weak agreement, ICC between .40 and .75 indicates moderate agreement, and ICC > .75 indicates strong agreement (Fleiss, Levin, & Paik, 2003). Other researchers have suggested that when analyzing the reliability of high stakes clinical measures, that the criteria of ICC > .90 should be used to measure acceptable reliability (Portney and Watkins, 2009). Given that the US-NSA is
generally used as a screening tool and based on precedent set in a previous NSA translation study (Lima et al., 2010), however, the less conservative methods of interpretation were used in the current study.

Results

Demographics

Nineteen stroke survivors participated in this study. One participant dropped out due to a busy schedule and another was excluded due to severe cognitive impairment, resulting in a final sample of 17 patient participants (8 males, 9 females). The average time post-stroke was 5.12 years. The average time between assessments was 4.11 days. Demographic data for patients is summarized in Table 2. Six rehabilitation therapists (one physical therapist, five occupational therapists) and one rehabilitation student participated in this study, for a total of seven assessor participants. Demographic data for assessors is summarized in Table 3.

Inter-rater Reliability

The inter-rater reliability of the US-NSA test items ranges from moderate to strong agreement (Table 4). Tactile Sensation (Light Touch, Pressure, Pinprick) for the upper extremity and for the lower extremity were found to have strong agreement between raters. Sharp-Dull Discrimination for the upper extremity indicated a strong level of agreement, although the confidence interval is fairly broad. Sharp-Dull Discrimination for lower extremity was found to have the lowest level of agreement out of all the test items. The authors found strong agreement for Proprioception for the upper and lower extremities. Stereognosis was also found to have the strongest agreement between raters. Reliability scores for Tactile Sensation (Light Touch, Pressure, and Pinprick), and Sharp-Dull Discrimination test items for the lower extremity were found to be lower than those for the upper extremity.

Perceptions of US-NSA

Therapists were asked for feedback about testing procedures and perceptions of their success implementing the measure. The authors did not keep records of the length of time to administer the measure, but all assessments were completed within a 30-minute session. As assessors administered the US-NSA more times, they reported that the speed of testing seemed to increase. In addition, they reported
feeling more comfortable and efficient administering the assessment after completing multiple trials. Therapists reported difficulty positioning some patients in the testing position (i.e., in supine with palms facing up, or “anatomical position”) due to the client’s limited range of motion, presence of contractures, and/or abnormal tone. Therapists expressed that they used their best clinical judgment in these situations. For example, with a patient with high tone, one assessor reported using tone management techniques and manual positioning prior to administering the US-NSA. With the same patient, the second assessor reported completing testing with the patient in his natural presenting position.

There appeared to be some confusion regarding the testing procedures for Tactile Sensation (Light Touch, Pressure, Pinprick) and Sharp-Dull Discrimination. If a patient scored 8/8 on Light Touch for a limb, then Pressure and Pinprick should be automatically assigned full scores. Sharp-Dull Discrimination should only be tested on a limb if the patient receives full scores (i.e., 8/8) on each Tactile Sensation subtest (Light Touch, Pressure, Pinprick) for that limb; otherwise Sharp-Dull Discrimination should not be administered and the patient automatically receives a 0/8 for this test item. Based on the way the score sheets were filled out and through conversations with the assessors, it appeared that they did not follow these instructions correctly. As such, there were several occasions when Pressure and Pinprick were tested when they should have been automatically assigned 8/8 for the limb, and when Sharp-Dull Discrimination was tested when instead it should have been assigned a 0/8. Occasionally, assessors added both limbs together, resulting in a score out of 16, rather than adding the limbs separately to achieve a score out of eight. Other times the assessors added the Stereognosis score incorrectly, marking a total score out of eight instead of out of 12. These scoring mistakes were reviewed and corrected before entering the data.

**Discussion**

This study analyzed the inter-rater reliability of a United States version of the Nottingham Sensory Assessment. The authors concluded that the US-NSA demonstrates strong inter-rater reliability for the majority of the subtests, indicating acceptability for use in clinical practice. Stereognosis, Light Touch, Pressure (upper extremity), Pinprick (upper extremity), and Proprioception were found to be the
most reliable. Sharp-Dull Discrimination, Pressure (lower extremity), and Pinprick (lower extremity) were found to be the least reliable.

The reliability coefficients of the US-NSA are generally lower ($r = .496$ to $r = .939$) than those found in the Brazilian version ($r = .97$ to $r = 1.00$) (Lima et al., 2010) and in the EmNSA ($K_w = .70$ to $K_w = 1.00$) (Stolk-Hornsveld et al., 2006). The differences in the sample demographics between the current study and these studies may explain the lower reliability findings in this study. The patient sample used in this study was smaller and had on average older participants compared to samples in previous NSA reliability studies (Connell, 2007; Gaubert and Mockett, 2000; Lima et al., 2010; Lincoln et al., 1991; Lincoln et al., 1998; Stolk-Hornsveld et al., 2006). Older adults are more likely to have secondary health issues, which may have added to more differences in sensory and physical abilities and therefore may have increased the challenges with testing this sample. Further, in a number of previous studies the sample populations were acute stroke patients (Connell, 2007; Gaubert and Mockett, 2000; Lincoln et al., 1991; Lincoln et al., 1998; Stolk-Hornsveld et al., 2006), whereas this study examined chronic stroke patients. This is significant because acute stroke patients are much less likely to have developed contractures or other secondary complications compared to chronic stroke patients. In the current study, the presence of contractures and abnormal tone appeared to affect some patients’ abilities to achieve the testing position. The assessors in the current study expressed uncertainty and differed in the way that they managed these positioning difficulties, which may have then contributed to decreased reliability.

This study is unique in that it employed seven assessors while most other studies employed only two assessors (Lima et al., 2010; Lincoln et al., 1991; Lincoln et al., 1998; Stolk-Hornsveld et al., 2006). The inclusion of five more assessors likely added more variability to the data because of individual differences between therapists, which may have contributed to decreased reliability. Using multiple assessors without a fully crossed design (i.e., having each assessor test every patient) necessitated the use of a one way random ICC (1, 1), which is a more conservative statistic that accounts for variance in the assessors as well as patients (Shrout & Fleiss, 1979). This will generally lead to lower reliability.
coefficients because of the need to account for variance in the assessors and patients (Shrout & Fleiss, 1979). The mistakes made by assessor participants likely contributed to lower reliability scores as well.

In the current study reliability coefficients for the lower extremity were generally lower than those of the upper extremity. Similar to these findings, Stolk-Hornsveld et al. (2006) found that lower extremity reliability scores were generally lower than those of the upper extremity in the EmNSA. This is not surprising as the upper extremity has greater representation in the brain than the lower extremity, suggesting that the upper extremity may have more opportunities for sensation perception than the lower extremity. This phenomenon may have contributed to more consistent patient responses during upper extremity testing and therefore higher rater agreement.

**Tactile Sensation**

Light Touch and Pressure had strong agreement for both extremities. Lima et al. (2010) and Stolk-Hornsveld et al. (2006) also found reliability coefficients for these subtests that indicate suitability for clinical use of their respective measures. In the US-NSA Pinprick was found to have strong agreement on the upper extremity, but only fair agreement for the lower extremity. This pattern of reliability has not been seen consistently in the inter-rater reliability of previous versions of the measure (Lincoln et al., 1991; Lincoln et al., 1998; Stolk-Hornsveld et al., 2006). In the current study, the agreement for Pinprick may have been lower compared to Light Touch and Pressure for a few reasons. One potential reason may be that assessors applied different amounts of pressure when using the pin compared to using the cotton ball or fingertip. The pin is a much narrower point than the cotton ball or fingertip, which may have lead to poorer detection and thus less consistent client responses.

**Sharp-Dull Discrimination**

Sharp-Dull Discrimination collectively had the lowest reliability of the subtests. In the EmNSA Sharp-Blunt Discrimination had lower reliability than Tactile Sensation (Light Touch, Pressure, and Pinprick) (Stolk-Hornsveld et al., 2006). The current authors hypothesize that this may have been due to assessor confusion regarding procedures to automatically assign scores. Based on analysis of errors, the
instructions and score sheet have been modified for increased clarity and ease of use. It is hypothesized that with these changes the reliability of this test may have been higher.

**Proprioception**

Interestingly, despite the overall trends of lower reliability found in this study compared to other studies, the reliability coefficients for Proprioception in the US-NSA were found to be higher than those found by Stolk-Hornsveld et al. (2006) for both upper and lower limbs. The US-NSA’s Proprioception test was closely based on the EmNSA method of proprioception (Stolk-Hornsveld et al., 2006). However, the current authors added more detail to the administration instructions to increase clarity during the modification period. The authors propose that these changes may have contributed to the higher reliability of the US-NSA version of this test.

**Stereognosis**

Stereognosis had strong agreement and was the subtest with the highest reliability in the current study. This suggests that the US-NSA version of stereognosis is reliable and appropriate for clinical use. Considering stereognosis reliability coefficients from previous studies: the ICC value in the US-NSA was lower than the range of ICCs for the Brazilian version of the NSA (Lima et al, 2010) and the kappa coefficients from the shortened rNSA (Connell, 2007), and was higher than the range of kappa coefficients found for the rNSA version (Gaubert & Mockett, 2000).

**Clinical Utility**

Based on assessor feedback and analysis of mistakes, some aspects of the score sheet and administration instructions have been modified to increase clarity and ease of scoring. The authors hypothesize that these additional changes to the testing materials will further enhance the inter-rater reliability of the US-NSA. In response to procedural and mathematical errors, the score sheet was revised to include a number of changes that the authors believe will improve the ease of use of the measure. For each test item, the authors added a total possible score in the totals rows to improve the simplicity and accuracy of scoring (Appendix A). The revised score sheet also includes reminders about the
administration procedures, specifically related to automatically assigning scores for the Pressure, Pinprick, and Sharp-Dull Discrimination tests.

The use of multiple assessors presented examples of how different therapists learn and administer this assessment, thus providing insight into how this measure may be approached in clinical practice. This process also gave the authors information about strengths and areas of improvement in the training, clarity of instructions, and ease of use of the US-NSA. After completing the training module for this study and with access to the authors, the assessors still made errors. This suggests a need for access to an effective training program with an emphasis on the unique procedures of the US-NSA to optimize clinician performance. The authors hypothesize that the following recommendations may increase therapists’ ability to accurately implement the US-NSA: offer access to training videos to demonstrate correct testing procedures, provide example score sheets to clarify how to complete automatic scoring for Tactile Sensation (Light Touch, Pressure, Pinprick) and Sharp-Dull Discrimination, and recommend a set number of times to administer and score the measure before use in clinic practice.

All evaluation sessions were completed within 30 minutes. There were a number of factors that may have contributed to increased testing time including transferring, donning and doffing orthoses, and changing clothes. The authors hypothesize that without these considerations, testing may have been completed more efficiently. However, these conditions mirror those found in clinical practice and contribute to generalizability. The assessors also reported that with practice they felt more efficient administering the US-NSA. Perhaps as assessors gain experience with the measure, they will be able to complete the US-NSA more quickly thereby increasing clinical utility.

Limitations

This research should be interpreted within the context of its limitations. There are a few limitations to this study. First, seven assessors of varying backgrounds and specialties participated in the present study to test inter-rater reliability. These assessors had limited time to learn and practice the new measure before the research trials began and they made some errors, which may have affected the reliability findings. The authors believe, however, that these limitations can also be viewed as strengths.
because these conditions mirror those often found in clinical practice. This adds to the generalizability of the current study and suggests that the US-NSA may be reliable in a real-world setting.

This study also differed from previous NSA reliability studies in terms of the inclusion and exclusion criteria (Gaubert & Mockett, 2000; Lima et al., 2010; Stolk-Hornsveld et al., 2006). Stroke survivors who had co-morbidities that may be other potential sources of sensory impairment (e.g. diabetic neuropathy) were not excluded from this study. The authors believed that this change would not affect the reliability of the measure, as the US-NSA is appropriate to use with patients regardless of co-morbidities. As a result, this study did not exclude stroke survivors with co-morbid sensory deficits, including one patient with spinal stenosis. This decision to include stroke survivors who had co-morbidities further represents the real-world clinical conditions under which this measure would be used.

Furthermore, a number of patients were unable to assume the standardized testing position (i.e., in supine with palms up) due to limited range of motion, presence of contractures, and/or abnormal tone. These patient characteristics, however, also closely mirror clinical conditions and therefore strengthen the generalizability of this study while still maintaining acceptable levels of reliability under these conditions. The authors also employed a convenience sample to recruit patient and assessor participants, which decreases the generalizability of this study to other clinics across the United States.

**Implications for OT**

Research has shown that occupational therapists perform and value information gained from sensory measures, but very few therapists use standardized multimodal measures (Doyle et al., 2013b). Standardized assessments are a vital part of informed clinical decision-making and evidence-based practice. In the United States there is currently no “gold standard” measure for evaluating sensory impairments after stroke. The creation of the US-NSA provides U.S. therapists with an opportunity to administer a reliable, standardized, clinically feasible assessment to quickly gather information about a patient’s sensory presentation.

This study has implications for increasing the efficiency and efficacy of the evaluation process as well as intervention planning. Sensory training in conjunction with motor training after stroke has been
shown to be more effective than motor training alone (Byl et al., 2003; Celnik, Hummel, Harris-Love, Wolk, & Cohen, 2007; Floel et al., 2004; Smania, Montagnana, Faccioli, Fiaschi, & Aglioti, 2003). Using the US-NSA in practice will allow therapists to learn about the nature of individuals’ sensory deficits in order to address these issues appropriately throughout the rehabilitation process and potentially enhance functional outcomes.

In addition this measure can be used to support occupational therapy research and demonstrate the outcomes of occupational therapy. The information gained from this measure will also enable therapists to gather evidence in practice to confidently monitor the effectiveness of their own interventions. Implementing the US-NSA in practice will allow consistent comparisons of client outcomes between therapists and between facilities.

**Future Directions**

This study uncovered many possibilities for future research. The reliability of the US-NSA should be tested with the revised score sheet and administration instructions. The validity of the US-NSA could be explored by comparing correlations to another standardized sensory test, for example the Fugl-Meyer sensory subscales. This study also raised questions about the impact of a patient’s cognitive and communication skills on the reliability of sensory measures. Future studies could explore this relationship to determine if a need for alternative evaluation methods should be employed with such patients.

Many patients tested in this study were unable to assume the starting position due to secondary complications. Positioning patients in supine often required transferring patients, which added time to the evaluation. Further research could explore how different positioning affects the reliability of scores. Using a less restrictive position might decrease testing time, which would increase clinical utility. Examining patterns and differences in sensation in the upper and lower extremities could be explored to determine the most effective and efficient ways of evaluating sensory impairments. Future research could also investigate if the upper and lower extremities can be evaluated separately. Other studies could analyze the sensitivity of the US-NSA and the effectiveness of its use with other diagnoses. In addition, future research could examine the correlation of US-NSA with an assessment of motor function, for
example the Chedoke Arm and Hand Activity Inventory, to explore the relationship between sensory and motor deficits.

**Conclusions**

The US-NSA has demonstrated strong inter-rater reliability for all test items, with the exception of the Sharp-Dull Discrimination item when tested on the lower extremity. These results have shown that the US-NSA is a promising, multimodal sensory assessment available for use by U.S. occupational therapists and other rehabilitation professionals to support intervention planning, clinical decision-making, functional outcomes, and quality of care. This standardized multimodal sensory measure can be used to screen for sensory deficits that may be impacting function, safety, and independence. The US-NSA Administration Manual, score sheets, and instructions for assembling a testing kit are freely available through the University of Puget Sound making this measure accessible to occupational therapists in the United States.
References


Table 1

*Patient Participant Demographics by Individual*

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>Sex</th>
<th>Year of Stroke</th>
<th>Years Post Stroke</th>
<th>More Affected Side</th>
<th>Days between Evaluations</th>
<th>First Assessor</th>
<th>Second Assessor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>77</td>
<td>M</td>
<td>2009</td>
<td>6</td>
<td>L</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>71</td>
<td>M</td>
<td>2009</td>
<td>6</td>
<td>L</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>75</td>
<td>F</td>
<td>2010</td>
<td>5</td>
<td>L</td>
<td>7</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>53</td>
<td>F</td>
<td>2010</td>
<td>5</td>
<td>R</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>69</td>
<td>F</td>
<td>2009</td>
<td>6</td>
<td>L</td>
<td>8</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>85</td>
<td>M</td>
<td>2012</td>
<td>3</td>
<td>L</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>70</td>
<td>F</td>
<td>2012</td>
<td>3</td>
<td>L</td>
<td>2</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>53</td>
<td>M</td>
<td>2008</td>
<td>7</td>
<td>L</td>
<td>2</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>62</td>
<td>F</td>
<td>2010</td>
<td>5</td>
<td>R</td>
<td>13</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>62</td>
<td>F</td>
<td>1998</td>
<td>17</td>
<td>R</td>
<td>2</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>60</td>
<td>F</td>
<td>2010</td>
<td>5</td>
<td>L</td>
<td>7</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>12</td>
<td>62</td>
<td>M</td>
<td>2014</td>
<td>1</td>
<td>L</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>72</td>
<td>M</td>
<td>2014</td>
<td>1</td>
<td>L</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>73</td>
<td>F</td>
<td>2006</td>
<td>9</td>
<td>L</td>
<td>7</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>15</td>
<td>72</td>
<td>F</td>
<td>2011</td>
<td>4</td>
<td>R</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>71</td>
<td>M</td>
<td>2014</td>
<td>1</td>
<td>R</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>53</td>
<td>M</td>
<td>2012</td>
<td>3</td>
<td>L</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Table 2

*Patient Participant Characteristics*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left Hemisphere Stroke (Right)</td>
<td>12 (5) N/A</td>
</tr>
<tr>
<td>Male (Female)</td>
<td>8 (9) N/A</td>
</tr>
<tr>
<td>Mean Age (SD)</td>
<td>67.09 (9.12) years 53-85 years</td>
</tr>
<tr>
<td>Mean Years since Stroke (SD)</td>
<td>5.12 (3.77) years 1-17 years</td>
</tr>
<tr>
<td>Mean Days between Trials (SD)</td>
<td>4.11 (3.22) days 1-13 days</td>
</tr>
</tbody>
</table>

*Note.* $N = 17$; SD = standard deviation
### Table 3

**Assessor Participant Characteristics**

<table>
<thead>
<tr>
<th>Therapist ID</th>
<th>Years in Practice</th>
<th>Area of Practice</th>
<th>Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>SNF (Geriatric)</td>
<td>PT</td>
</tr>
<tr>
<td>2</td>
<td>8.5</td>
<td>SNF</td>
<td>OT</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>Adult Physical Rehabilitation</td>
<td>OT</td>
</tr>
<tr>
<td>4</td>
<td>34</td>
<td>SNF</td>
<td>OT</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>Acute Adult Rehabilitation</td>
<td>OT</td>
</tr>
<tr>
<td>6</td>
<td>39</td>
<td>Adult Physical Rehabilitation</td>
<td>OT</td>
</tr>
<tr>
<td>7</td>
<td>N/A</td>
<td>Student</td>
<td>OT</td>
</tr>
</tbody>
</table>

*Note. N = 7; SNF = skilled nursing facility*
Table 4

*Intraclass Correlation Coefficients (ICC) for Inter-rater Reliability on the US-NSA*

<table>
<thead>
<tr>
<th>Subtest</th>
<th>ICC</th>
<th>Confidence Interval (CI) 95%</th>
<th>p-values</th>
<th>Level of Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light Touch UE</td>
<td>0.920</td>
<td>0.799-0.970</td>
<td>&lt;.001</td>
<td>Strong</td>
</tr>
<tr>
<td>Pressure UE</td>
<td>0.848</td>
<td>0.637-0.942</td>
<td>&lt;.001</td>
<td>Strong</td>
</tr>
<tr>
<td>Pinprick UE</td>
<td>0.914</td>
<td>0.784-0.968</td>
<td>&lt;.001</td>
<td>Strong</td>
</tr>
<tr>
<td>Sharp-Dull Discrimination UE</td>
<td>0.818</td>
<td>0.575-0.929</td>
<td>&lt;.001</td>
<td>Strong</td>
</tr>
<tr>
<td>Proprioception UE</td>
<td>0.846</td>
<td>0.633-0.941</td>
<td>&lt;.001</td>
<td>Strong</td>
</tr>
<tr>
<td>Stereognosis</td>
<td>0.939</td>
<td>0.843-0.977</td>
<td>&lt;.001</td>
<td>Strong</td>
</tr>
<tr>
<td>Light Touch LE</td>
<td>0.858</td>
<td>0.657-0.945</td>
<td>&lt;.001</td>
<td>Strong</td>
</tr>
<tr>
<td>Pressure LE</td>
<td>0.809</td>
<td>0.556-0.926</td>
<td>&lt;.001</td>
<td>Strong</td>
</tr>
<tr>
<td>Pinprick LE</td>
<td>0.772</td>
<td>0.485-0.910</td>
<td>&lt;.001</td>
<td>Strong</td>
</tr>
<tr>
<td>Sharp-Dull Discrimination LE</td>
<td>0.496</td>
<td>0.048-0.781</td>
<td>0.16</td>
<td>Moderate</td>
</tr>
<tr>
<td>Proprioception LE</td>
<td>0.840</td>
<td>0.621-0.939</td>
<td>&lt;.001</td>
<td>Strong</td>
</tr>
</tbody>
</table>

*Note. N = 17; UE = upper extremity, LE = lower extremity; ICC < 4.0 indicates weak agreement, 0.40 – 0.75 indicates moderate agreement, > 0.75 indicates strong agreement.*
# US-NSA Score Sheet

**Pt Name:** __________________________  **Examiner:** ____________________________  
**Date:** ____________________________  **Dx:** ____________________________

**Side of body affected:** RIGHT / LEFT / BOTH / NEITHER

<table>
<thead>
<tr>
<th>Tactile Sensation</th>
<th>Sharp-Dull Discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body Sections</td>
<td>Light Touch</td>
</tr>
<tr>
<td></td>
<td>L</td>
</tr>
<tr>
<td>Fingers</td>
<td></td>
</tr>
<tr>
<td>Hand</td>
<td></td>
</tr>
<tr>
<td>Forearm</td>
<td></td>
</tr>
<tr>
<td>Upper Arm</td>
<td></td>
</tr>
<tr>
<td>Total Score UE</td>
<td>/8</td>
</tr>
<tr>
<td>Toes</td>
<td></td>
</tr>
<tr>
<td>Foot</td>
<td></td>
</tr>
<tr>
<td>Leg</td>
<td></td>
</tr>
<tr>
<td>Thigh</td>
<td></td>
</tr>
<tr>
<td>Total Score LE</td>
<td>/8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Score</th>
<th><strong>0</strong></th>
<th><strong>1</strong></th>
<th><strong>2</strong></th>
<th><strong>X</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABSENT</td>
<td>IMPAIRED</td>
<td>NORMAL</td>
<td>UNABLE TO TEST</td>
</tr>
</tbody>
</table>

1If pt receives 8/8 on UE for light touch, automatically assign 8/8 for UE pressure and pinprick. The same applies to LE.

2Only test sharp-dull if pt receives 8/8 on each of the tactile sensations for a limb.

---

*Occupational Therapy, University of Puget Sound 2015*

*Modified from the Revised Nottingham Sensory Assessment and Exumas MC Modifications to the (Revised) Nottingham Sensory Assessment*
**US-NSA SCORE SHEET**

<table>
<thead>
<tr>
<th>Joints</th>
<th>RIGHT</th>
<th>LEFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finger</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wrist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elbow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shoulder</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Score UE</td>
<td>/8</td>
<td>/8</td>
</tr>
<tr>
<td>Toe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ankle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knee</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Score LE</td>
<td>/8</td>
<td>/8</td>
</tr>
</tbody>
</table>

0 = ABSENT  1 = IMPAIRED  2 = NORMAL  X = UNABLE TO TEST

**STEREOGNOSIS**

- Comb
- Sponge
- Quarter (coin)
- Pencil
- Glass
- Washcloth (towel)

/12 Total Score

**Patient may identify quarter as ‘coin’ and washcloth as ‘towel’ to receive a score of 2**

0 = ABSENT  1 = IMPAIRED  2 = NORMAL  X = UNABLE TO TEST

**COMMENTS:**

---

*Occupational Therapy, University of Puget Sound 2015*

Modified from the Revised Nottingham Sensory Assessment and Exarms MC Modifications to the (Revised) Nottingham Sensory Assessment
US-NSA SCRIPT AND SCORING CRITERIA

Opening statement: “I’m going to look at the sensation in your arm and leg today.”

NOTE: If testing both sides, test the more affected side first.

Tactile Sensation

Light Touch
“First I’m going to touch you with this cotton ball.” Touch patient with cotton ball. “And I want you to say ‘yes’ when you feel it. I’m going to put this blindfold on you while we do the test.” Put blindfold on patient. “Ready?”

NOTE: If a score of 2 is assigned for all of a limb for light touch, the patient automatically receives a score of 2 for all the pressure and pinprick items for that limb.

Pressure
“Now I’m going to do the same thing, but this time I’ll touch you with my finger.” Touch patient with finger. “And I want you to say ‘yes’ when you feel it. I’m going to put the blindfold on you while we do the test.” Put blindfold on patient. “Ready?”

Pinprick
“Now we’re going to do one more similar test. This time I’m going to use something a little bit sharp, and you’ll feel a slight poke.” Touch patient with sharp end of Neurotip. “And I want you to say ‘yes’ when you feel it. I’m going to put the blindfold on you while we do the test.” Put blindfold on patient. “Ready?”

Scoring Criteria

0 Absent Patient fails to identify the test sensation on all three occasions
1 Impaired Patient identifies the test sensation on only one or two occasions
2 Normal Patient identifies the test sensation on all three occasions
X Unable to test

Sharp-Dull Discrimination

NOTE: Only tested if patient receives 8/8 on each of the tactile sensations (light touch, pressure, and pinprick) for a limb.

“Now I’ll either touch you with this pin or I’ll touch you with my finger. I want you to say ‘sharp’ when I touch you with this.” Touch with sharp end of Neurotip. “Or ‘dull’ when I touch you with my finger.” Touch with finger. “I’m going to put the blindfold back on.” Put blindfold on patient. “Ready?”

Scoring Criteria

0 Absent Patient fails to correctly identify the test sensation on all six occasions
1 Impaired Patient correctly identifies the test sensation, but on less than six occasions
2 Normal Patient correctly identifies the test sensation on all six occasions
X Unable to test
US-NSA SCRIPT AND SCORING CRITERIA

Proprioception

"Now I am going to test your ability to sense movement. I'm going to move a joint and ask you to tell me which direction it moved. Here's an example. " Demonstrate both thumb movements without blindfolding the patient. "So I'll move your thumb like this and ask you to tell me, 'Is your thumb being bent or straightened?' If you can't tell which way it moved, that's ok just let me know. Ready?"

**If patient is unsure of the direction of the movement ask: "Did you feel it move?"**

Fingers: "Is your thumb being bent or straightened?"

Wrist: "Is your hand moving up or moving down?"

Elbow: "Is your elbow being bent or straightened?"

Shoulder: "Is your arm moving towards you or away from you?"

Toes: "Is your toe moving up or moving down?"

Ankle: "Is your foot moving up or moving down?"

Knee: "Is your knee being bent or straightened?"

Hip: "Is your thigh moving towards you or away from you?"

Scoring Criteria

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Absent</td>
</tr>
<tr>
<td>1</td>
<td>Impaired</td>
</tr>
<tr>
<td>2</td>
<td>Normal</td>
</tr>
<tr>
<td>X</td>
<td>Unable to test</td>
</tr>
</tbody>
</table>

Stereognosis

NOTE: Do NOT show patient the objects beforehand.

"For the next test, I'm going to put some objects in your hand. I want you to tell me what the object is just by feeling it. I'm going to put the blindfold back on for this test. " Put blindfold on patient. Do NOT give positive or negative feedback after the patient responds.

**If patient cannot name the object, ask, "Can you describe what it feels like?"**

Scoring Criteria

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Absent</td>
</tr>
<tr>
<td>1</td>
<td>Impaired</td>
</tr>
<tr>
<td>2</td>
<td>Normal</td>
</tr>
<tr>
<td>X</td>
<td>Unable to test</td>
</tr>
</tbody>
</table>
Appendix B

UNIVERSITY OF NOTTINGHAM SENSORY
ASSESSMENT (US-NSA) ADMINISTRATION MANUAL

I. GENERAL INSTRUCTIONS

The US-NSA includes four tests: Tactile Sensation, Sharp-Dull Discrimination, Proprioception, and Sterognosis. Tactile Sensation includes three subtests: light touch, pressure, and pinprick. During testing, ideally the patient will be dressed in loose clothing with defined points of contact easily accessible. During all tests the patient will lie in supine with forearms in supination while wearing a blindfold. Each test should be explained prior to administration and, if necessary, will be demonstrated on the less affected side. Testing should be performed distally to proximally in the order indicated on the score sheet. This measure can be used as a screening tool if the examiner tests only the more affected side. It can also be used as a more comprehensive assessment if the examiner wishes to test both sides of the body. Within the Tactile Sensation and Sharp-Dull Discrimination tests, each body section (e.g., fingers, hand, forearm, upper arm, toes, foot, leg, thigh) should be tested only once at the three defined points of contact. Administer stimuli for 1-2 seconds. No more than 2-5 seconds should pass between administrations of stimuli for each test item. If a body section is unable to be tested (e.g., open wound, casting, hypersensitivity, loss of limb, etc.) assign an X. Be sure to note any deviations from standard procedure in the comments section on the score sheet.

II. TACTILE SENSATION

There are three Tactile Sensation subtests: light touch, pressure, and pinprick. The eight body sections to be tested are fingers, hand, forearm, upper arm, toes, foot, leg, and thigh. Each body section has three defined points of contact, which are described in detail below. During each subtest, each body section is stimulated three times in a random order, once at each defined point of contact. The patient must indicate verbally or non-verbally (in a manner previously agreed upon by the patient and examiner) when he/she feels the test stimulus.

**Light touch**: touch the skin, once at each defined point of contact, lightly with a cotton ball

**Pressure**: apply pressure to the skin sufficient to just deform the skin contour, using the examiner’s index finger, once at each defined point of contact

**Pinprick**: prick the skin using a Neurotip™ sufficient to just deform the skin contour, once at each defined point of contact

Scoring criteria for light touch, pressure, and pinprick

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Absent</td>
</tr>
<tr>
<td>1</td>
<td>Impaired</td>
</tr>
<tr>
<td>2</td>
<td>Normal</td>
</tr>
<tr>
<td>X</td>
<td>Unable to test</td>
</tr>
</tbody>
</table>

**NOTE**: If a score of 2 is assigned for all of a limb for light touch, the patient automatically receives a score of 2 for all the pressure and pinprick items for that limb.

---

1 Occupational Therapy, University of Puget Sound 2015

Modified from the Revised Nottingham Sensory Assessment and Erasmus MC Modifications to the Revised Nottingham Sensory Assessment
DEFINED POINTS OF CONTACT

List of defined points of contact for stimulating tactile sensations (light touch, pressure, pinprick) and sharp-dull discrimination. A diagram of these points of contact is provided below and on the US-NSA score sheet.

1. Fingers
   a. Distal phalanx of 5th digit, palmar aspect
   b. Distal phalanx of 3rd digit, palmar aspect
   c. Distal phalanx of 1st digit, palmar aspect

2. Hand
   a. 2nd metacarpal, distal palmar aspect
   b. 5th metacarpal, distal palmar aspect
   c. Center of thenar eminence

3. Forearm
   a. Ulnar styloid, anterior aspect
   b. Center of forearm, anterior aspect
   c. 2 cm distal to elbow joint line, anterolateral aspect

4. Upper Arm
   a. 2 cm proximal to the elbow joint, anteromedial aspect
   b. Center of anterior aspect of the humerus
   c. 2 cm distal to the acromion, lateral aspect

5. Toes
   a. Distal phalanx of 5th digit, plantar aspect
   b. Distal phalanx of 3rd digit, plantar aspect
   c. Distal phalanx of 1st digit, plantar aspect

6. Foot
   a. Base of 5th metatarsal bone, dorsal aspect
   b. 2nd metatarsal, distal dorsal aspect
   c. Center of midtarsal line, dorsal aspect

7. Leg
   a. Medial malleolus, medial aspect
   b. Center of the anterior border of the tibia
   c. Fibular head, lateral aspect

8. Thigh
   a. Medial femoral epicondyle, medial aspect
   b. Center of line of femur, anterior aspect
   c. Greater trochanter
IV. SHARP-DULL DISCRIMINATION

NOTE: Only tested if patient receives 8/8 on each of the tactile sensations (light touch, pressure, and pinprick) for a limb.

All eight body sections are to be tested (fingers, hand, forearm, upper arm, toes, foot, leg, and thigh). Each body section has three defined points of contact using the defined points of contact listed above. Stimulate the skin a total of six times in a random order once at each defined point of contact with each stimuli (i.e., three times with the sharp end of the Neurotip™ and three with the index finger). The patient must indicate verbally or non-verbally (in a manner previously agreed upon by the patient and examiner) if the test sensation is sharp or dull.

Scoring criteria for Sharp-Dull Discrimination
0 Absent Patient fails to correctly identify the test sensation on all six occasions
1 Impaired Patient correctly identifies the test sensation, but on less than six occasions
2 Normal Patient correctly identifies the test sensation on all six occasions
X Unable to test

V. PROPRIOCEPTION

This tests measures kinesthesia (awareness of movement) and selected proprioceptive discrimination (discrimination of movement direction). Each joint is tested three times, one joint at a time, with a specified passive movement. (Full description of specific starting positions, movements, and the examiner’s hand grips are described below).

Large joints (shoulder, elbow, hip, and knee) should be moved through about 25% of their full passive range of motion (PROM). Small joints (fingers, wrist, toes, and ankle) should be moved through their full available PROM. The examiner may demonstrate the procedure by providing three different practice movements with the patient’s eyes open. Each joint is then moved three times. Return to starting position between each trial. The patient is asked to indicate verbally or non-verbally the direction of the movement taking place. If the patient is unable to identify the direction, he/she is instead asked to identify when the movement is taking place. In that case, the highest score he/she can receive is a 1.

Scoring criteria Proprioception
0 Absent Patient does not detect the movement taking place
1 Impaired Patient detects the movement taking place, but the direction is not correct on all three occasions
2 Normal Patient correctly detects the direction of the movement taking place on all three occasions
X Unable to test
DESCRIBED JOINT MOVEMENTS FOR PROPRIOCEPTION

Starting Position: Patient will be lying in supine with forearm in supination with a blindfold on. Large joints (shoulder, elbow, hip, and knee) should be moved through about 25% of their full PROM, while small joints (fingers, wrist, toes, and ankle) are moved through their full available PROM.

Fingers
Movement: flexion and extension of the distal phalanx of the thumb
Examiner Hand Grips:
- The moving hand grasps the distal phalanx of patient’s thumb, with the examiner’s thumb and index finger on the lateral and medial aspect of the patient’s thumb
- The fixing hand grasps the proximal phalanx of patient’s thumb using the examiner’s thumb and index finger
Ask the patient: “Is your thumb being bent or straightened?”

Wrist
Movement: flexion and extension of the wrist; place elbow in starting position of 20°-30° flexion
Examiner Hand Grips:
- The moving hand grasps the patient’s hand, with the examiner’s thumb and index finger on the lateral and medial aspect of the hand
- The fixing hand grasps the distal end of the forearm, with the examiner’s thumb and index finger on the lateral and medial aspect
Ask the patient: “Is your hand moving up or moving down?”

Elbow
Movement: flexion and extension of the elbow; place elbow in starting position of 90° flexion
Examiner Hand Grips:
- The moving hand grasps the distal end of the forearm, with the examiner’s thumb and fingers on the anterior and posterior aspect of the forearm
- The fixing hand grasps the distal end of the humerus
Ask the patient: “Is your elbow being bent or straightened?”

Shoulder
Movement: abduction and adduction of the shoulder; place elbow in 90° flexion, place the shoulder in middle of available passive range; lift the arm enough to allow the movements to occur without sliding against surface
Examiner Hand Grips:
- The 1st guiding hand grasps the distal end of the forearm, with the examiner’s thumb and fingers on the anterior and posterior aspect of the forearm
- The 2nd guiding hand cups the flexed elbow
Ask the patient: “Is your arm moving towards you or away from you?”
Toes
Movement: flexion and extension of the 1st metatarsophalangeal joint (big toe)

Examiner Hand Grips:
• The moving hand grasps the distal phalanx of patient’s big toe, with examiner’s thumb and index finger on the lateral and medial aspect of the big toe
• The fixing hand grasps the 1st metatarsal bone, just proximal to the metatarsophalangeal joint, with examiner’s thumb lateral and index finger medial

Ask the patient: “Is your toe moving up or moving down?”

Ankle
Movement: flexion and extension of the ankle joint

Examiner Hand Grips:
• The moving hand grasps the patient’s foot, with examiner’s thumb on lateral margin and fingers on the medial margin of the foot
• The fixing hand grasps the distal end of the tibia and fibula

Ask the patient: “Is your foot moving up or moving down?”

Knee
Movement: flexion and extension of the knee with hip and knee joints in 90° flexion

Examiner Hand Grips:
• The moving hand grasps the calcaneus, with examiner’s thumb on the medial aspect and the examiner’s fingers cupped inferiorly; in this position, the patient’s foot should be supported by the examiner’s lower forearm
• The fixing hand grasps the distal end of the femur, with the examiner’s thumb and fingers on the lateral and medial aspect of the femur

Ask the patient: “Is your knee being bent or straightened?”

Hips
Movement: flexion and extension of the hip joint, starting with the hip and knee joints in 90° flexion, maintain knee joint position as you move the hip

Examiner Hand Grips:
• The guiding hand grasps the calcaneus, with examiner’s thumb on the medial aspect and the examiner’s fingers cupped inferiorly; in this position, the patient’s foot should be supported by the examiner’s lower forearm
• The moving hand grasps the distal end of the femur, with the examiner’s thumb and fingers on the lateral and medial aspect of the femur

Ask the patient: “Is your thigh moving towards you or away from you?”
VI. STEREOGNOSIS

A total of six objects will be used in this test: quarter, pencil, comb, sponge, washcloth, glass. Do not show the objects to the patient before testing. Each object will be placed, one at a time, in the patient’s hand for up to 30 seconds. The patient can manipulate the object with one hand as needed. The patient must identify or describe the item verbally, or pair-match (i.e., select a match from a second set of items) the item non-verbally. Patient may say “coin” for quarter and “towel” for washcloth to receive a score of 2. If testing both sides of the body, the more affected side should be tested first. The object may be moved around the affected hand by the examiner if the patient is unable to manipulate the object. Do not provide any verbal cues. Make a note in the comments section of the score sheet if patient uses pair-matching or is non-English speaking, or if therapist manipulated any objects for the patient.

Scoring criteria for Stereognosis
0 Absent Patient is unable to identify the object in any manner
1 Impaired Patient correctly identifies some features of the object
2 Normal Patient is able to correctly name or match the object
X Unable to test
<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
</table>
| Blindfold/sleep mask  | Dimensions: 0.5" L x 4.13" W x 9.25" H  
Material: Fleece, Nylon  
Strap around head, not around ears | Travel Smart Sleep Mask Black (available at Target©)            |
| Neurotip™             | Sterile, single-use neurological examination pins                                               | Neurotips Testing Pins, 100/Bx                                  |
| Cotton ball           | 100% cotton, 100 balls per bag                                                                  | Up & Up Jumbo Cotton Balls 200 ct (available at Target©)        |
| Coin                  | Quarter (United States coin) – regular issue design  
Other designs acceptable (e.g., bicentennial, statehood) | United States regular issue quarter                             |
| Pencil                | Material: Wooden pencil  
Yellow, hexagon-shaped barrel  
Unsharpened #2  
Intact eraser  
PMA certified nontoxic | Dixon Ticonderoga #2 Pencils with Eraser                                                           |
| Comb                  | Material: Plastic  
Size: 5 inch  
Pocket comb  
Fine teeth OR Dual fine and wide teeth | Conair Pocket Comb  
Goody 5" Pocket Comb                                                |
| Sponge                | Multipurpose sponge  
Material: Cellulose  
Dimensions: 7.7" L x 4.2" W  
Color may vary  
No scrubbed (rough) siding | O-Cel-O™ 2 Pk. Sponge                                                                 |
| Washcloth             | Weave Type: Terry  
Material: 100% Cotton  
Dimensions: 12.0" L x 12.0" W  
Fabric Weight: 0.11 Lb. | 8 Pack Cotton Washcloths White (available at Target©)          |
| Glass                 | Capacity (volume): 12-16 Oz.  
Material: Glass  
Glassware Style: Tumbler  
Smooth edges (no bumps or ridges) | Libbey Clear Flare Tumblers (available at Target©)              |
Permission for Scholarly Use of Thesis

We hereby give permission for the University of Puget Sound and the Occupational Therapy Program to hold copyright and make this master's research project available for scholarly purposes in perpetuity. We understand that a copy will be housed in the Collins Memorial Library, another in the Resource Room of the School of Occupational Therapy & Physical Therapy, and, if appropriate, a copy may be placed on the University's web server for access by the public. We further understand that, if we submit our project for publication and the publisher requires the transfer of copyright privileges, the University of Puget Sound will relinquish copyright, and remove the project from its website if required by the publisher.

Name: Amy Kaestner, OTS

______________________________  Date: ________________
Signature of MSOT Student

Name: Rebecca Miller, OTS

______________________________  Date: ________________
Signature of MSOT Student