Symptom Validity Testing (SVT): General Update on Current Practice Parameters and Specific Application in a Clinical Population

Daniel L. Drane, Ph.D., ABPP (CN)
Assistant Professor of Neurology
Director, Neuropsychology Laboratory
UW Regional Epilepsy Center

Symptom Validity Testing (SVT) - refers to the development of measures and procedures that attempt to assess a test taker's level of effort.

SVT is intended to determine if someone is optimally engaged in the testing process and does not directly assess the etiology for poor performances.

Failed SVT suggests that test data is invalid but does not necessarily equate to a diagnosis of malingering.

SVT is based on the underlying assumption that test performance not only reflects brain functioning but also a whole host of intervening variables (e.g., motivation, alertness, compliance).

Possible Reasons for SVT Failure:

- True brain dysfunction (severe in nature)
- Disinterest in being assessed
- Somatizational or hysterical features
- Psychiatric/emotional symptomatology
- Malingering
- Transient physiological factors
  - Severe sleep deprivation
  - Acute pain of a severe nature
  - over-sedation from medications

Comment on Classification Accuracy on SVT Measures

- False Positive – Someone who scored within the fail range on an SVT measure who was not functioning in an independent fashion (e.g., patients in group homes with caregivers who do not manage their own finances).
- False Negative – Someone who passed an SVT measure despite clearly exhibiting evidence of poor effort on other tests.
- True Negative – Someone who scored within the pass range on an SVT measure who also did not demonstrate any evidence of poor effort on other tests.
- True Positive – Someone who scored within the fail range of an SVT measure who was functioning independently (e.g., living independently, working, managing their own finances) and who may or may not have shown clear-cut patterns of suspect effort on other measures.

Current Status of SVT in Neuropsychological Assessment

-Controversy remains over the need for SVT, the validity of SVT assessment, and the proper means of dealing with invalid performances.

-Nevertheless, there has been a proliferation of SVT measures and procedures.

-This has lead to initial efforts to develop standardized procedures for SVT use (see NAN position paper)

"The clinician should be prepared to justify a decision not to assess symptom validity as part of a neuropsychological assessment."

Bush, S. S., Ruff, R. M., Troster, A. I., et al. (2005). Symptom validity assessment: Practice issues and medical necessity. *Archives of Clinical Neuropsychology*, *20*, 419-426.

Methods of SVT Assessment

- Consistency (of symptoms)
- Performance on neurocognitive tests
- Performance on psychological tests
- Symptom validity tests
- Forced-choice tests

From NAN Position Paper

Consistency of symptoms

- Self-reported history that is inconsistent with documented history
- Self-reported symptoms that are inconsistent with known patterns of brain functioning.
- Self-reported symptoms that are inconsistent with behavioral observations.
- Self-reported symptoms that are inconsistent with information obtained from reliable collateral informants.
- Self-reported presence or absence of symptoms that are inconsistent with performance levels on psychometric tests.

From NAN Position Paper

Performance on Neurocognitive Tests

- Performance consistent with feigning on empirically derived indices obtained from scores of ability measures.
- Performance patterns on ability meaures indicative of invalid responding.
- Inconsistencies between test results and known patterns of brain functioning.
- Inconsistencies between test results and observed behavior.

From NAN Position Paper

Performance on Neurocognitive Tests

- Inconsistencies between test results and reliable collateral reports.
- Inconsistency between test results and documented background information.

From NAN Position Paper

Performance on Psychological Tests

- Evidence of exaggerated or fabricated problems may be evident from the original and more recently developed validity scales of self-report psychological tests, such as the MMPI-2.

From NAN Position Paper

Performance on SVT Measures

- Performance below established cut-off scores on one or more well-validated tests designed to measure exaggeration or fabrication of cognitive deficits suggests insufficient effort to do well.

From NAN Position Paper

Performance on Forced-Choice Tests

- Performance on one or more forced-choice measures of cognitive functioning that falls below chance to a statistically significant degree indicates biased responding.

From NAN Position Paper

Suggestions Regarding SVT Procedures (selected)

- Remain abreast of trends in the SVT literature.
- Use a multi-method approach.
- Inform the examinee at the outset of the evaluation and as needed during the evaluation that good effort and honesty will be required.
- Disperse SVTs or measures with SV indicators throughout the evaluation, with administration of at least one SVT early in the assessment.
- Report the results of SVT performance.

From NAN Position Paper

Suggestions Regarding SVT Interpretation (selected)

- Give greater weight to the results of SVT measures than to subjective indicators of suboptimal effort.
- Invalid performance on a measure of personality does not allow for an a priori conclusion that neurocognitive test results are also unreliable, and vice versa.
- When evidence of invalid performance exists, scores on cognitive ability tests may be interpreted as representing the examinee's minimum level of ability.

From NAN Position Paper

Suggestions Regarding SVT Interpretation (selected)

- Strong evidence of invalid performance on SVTs or other indicators of symptom validity raise doubt about the validity of all neurocognitive test results. In the presence of invalid performance on measures or indices of SV, interpretations of performances on other tests as valid would need to be justified.

From NAN Position Paper

Suggestions Regarding SVT Interpretation (selected)

- Performance slightly below cut-off on one SVT may not justify an interpretation of biased responding; converging evidence from additional indicators may be required.
- If an evaluation that has been discontinued due to insufficient effort or invalid responding is later continued, the confidence that could be placed in the validity of the results would remain limited.

From NAN Position Paper

Suggestions Regarding SVT Interpretation (selected)

- The examinee's cultural background should be evaluated.
- Cultural factors may lead to either exaggeration of symptoms or denial of symptoms without any conscious or unconscious motivation to "deceive."
- Simply because an SVT has been validated in the majority culture, it does not mean that the test is equally valid with individuals from a minority culture.

From NAN Position Paper

Frequency of Use of SVT Measures

	Percent Responding			
	Never	Rarely	Often	Always
Test of Memory Malingering (TOMM)	29.2	25.0	20.8	25.0
Rey 15-Item Test	25.0	37.5	20.8	12.5
Recognition Memory Test (RMT)	50.0	25.0	8.3	16.7
Word Memory Test (WMT)	50.0	29.2	16.7	4.2
Validity Indicator Profile (VIP)	66.7	12.5	12.5	8.3
Computerized Assessment of Response Bias	66.7	16.7	8.3	8.3
Portland Digit Recognition Test (PDRT)	58.3	25.0	8.3	8.3
Victoria Symptom Validity Test (VSVT)	79.2	4.2	8.3	8.3
Digit Memory Test (DMT)	79.2	8.3	4.2	8.3

From: Slick, D. J., Tan, J. E., Strauss, E. H., & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Neuropsychology*, *19*, 465-473.

While there was considerable diversity in the choice of SVT, most experts (79%) reported using at least one specialized technique in every examination.

From: Slick, D. J., Tan, J. E., Strauss, E. H., & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Neuropsychology*, *19*, 465-473.

Many experts routinely evaluate effort using indexes from standard neuropsychological tests:

This includes the use of the Rey Auditory Verbal Learning Test, the California Verbal Learning Test, Digit Span (Reliable Digit Span), Wechsler Tests, Wisconsin Card Sorting Test, Category Test, and MMPI.

From: Slick, D. J., Tan, J. E., Strauss, E. H., & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Neuropsychology*, *19*, 465-473.

Is Cognitive Impairment Equal in Patients with Epileptic and Psychogenic Nonepileptic Seizures?

Daniel L. Drane, Ph.D.,[1] David J. Williamson, Ph.D.,[2-3] Elizabeth S. Stroup, Ph.D.,[1] Mark D. Holmes, MD[1], Matthew Jung, BS,[1] Erich Koerner, BS,[1] Alan J. Wilensky, MD, Ph.D.,[1]
& John W. Miller, MD, Ph.D.[1]

[1]Department of Neurology, Regional Epilepsy Center, University of Washington School of Medicine, Seattle, Washington
[2]Medical Affairs, Johnson & Johnson, Mobile, Alabama

[3]Department of Neurology, University of South Alabama

Several studies have concluded that patients with psychogenic nonepileptic seizures (PNES) experience cognitive deficits on neuropsychological testing as severe or worse than those experienced by patients with genuine epileptic seizures (ES) (1-5).

1. Dodrill CB, Holmes MD. Psychological and neuropsychological evaluation of the patient with non-epileptic seizures. In: Gates JR, Rowan AJ, eds. *Non-epileptic seizures*, 2nd ed. Boston: Butterworth-Heinemann, 2000:169-181.
2. Drake ME, Huber SJ, Pakalnis A, Phillips BB. Neuropsychological and event-related potential correlates of nonepileptic seizures. *Journal of Neuropsychiatry & Clinical Neurosciences* 1993;*5*:102-104.
3. Hermann BP. Neuropsychological assessment in the diagnosis of non-epileptic seizures. In: Gates JR, Rowan AJ, eds. *Non-epileptic seizures*. Boston: Butterworth-Heinemann, 1993:221-232.
4. Wilkus RJ, Dodrill CB, Thompson, PM. Intensive EEG monitoring and psychological studies of patients with pseudoepileptic seizures. *Epilepsia* 1984;*25*:100-107.
5. Wilkus RJ, Dodrill CB. Factors affecting the outcome of MMPI and neuropsychological assessments of psychogenic and epileptic seizure patients. *Epilepsia* 1989;*30*:339-347.

Several studies (1, 4, 5), for example, reported both groups perform outside of normal limits on approximately one half of the measures in a battery of neuropsychological tests.

Researchers have hypothesized that the severity of neurocognitive deficits in patients with PNES is due to their other medical difficulties, as they frequently report more neurological injury or disease than patients with epilepsy (e.g., head trauma, CNS infection, possible birth traumas). Such histories of neurologic insult are typically based upon self-report rather than objective data, however, and are rarely verified.

An alternative hypothesis to account for the apparent neurocognitive dysfunction of patients with PNES is that their poor scores result from inconsistent effort rather than true brain impairment.

Binder et al. (1998) found that, whereas patients with PNES and patients with ES exhibited equivalent neurocognitive dysfunction, the PNES group performed significantly worse on a SVT (the Portland Digit Recognition Test). Likewise, the neurocognitive performance of the PNES group was more strongly associated with SVT performance than was the ES group.

Binder LM, Kindermann SS, Heaton RK, Salinsky MC. Neuropsychologic impairment in patients with nonepileptic seizures. *Archives of Clinical Neuropsychology* 1998;*13*:513-522.

We predicted that:

(a) Patients with PNES would fail SVT at a higher rate than those with ES, on a well-validated neurocognitive battery sensitive to deficits seen with seizure disorders.

(b) Patients with PNES who pass SVT would significantly outperform both patients with ES and those with PNES who fail SVT, and

(c) Patients with PNES would report significantly more unverifiable neurological diseases or injuries than patients with epilepsy.

We classified patients on the basis of their ictal video EEG recordings as experiencing:

(a) Epileptic Seizures (ES: n = 70) – Evidence of definite ictal EEG abnormalities;
(b) Psychogenic Non-Epileptic Seizures (PNES: n = 43) – Episodes of unresponsiveness or behavioral abnormality in the absence of EEG changes;
(c) Indeterminate Spells (IS: n = 44) – No spells during monitoring or subjective feelings only, in the absence of EEG abnormality, unresponsiveness, or behavioral abnormality;

Patient Classification (Continued):

(d) Co-Occurrence Group (COG: n = 6) – Evidence of episodes fitting the criteria for both ES and PNES during the same or across multiple monitoring sessions; or
(e) Non-Epileptic Seizures of Other Origin (NESO: n = 3) – This included patients with spells resulting from medical conditions other than epilepsy (e.g., syncopal episodes).

Given the small size of the COG and NESO groups, we felt we would be unable to draw meaningful conclusions about them. Thus, we removed them from further analyses.

Likewise, we excluded those patients (n = 26) who experienced electrographically-confirmed seizure activity during any portion of the testing or within the 24-hour period preceding the testing, as data suggest that postictal patients may perform below the level typical of their interictal functioning.

Table 1. Comparison of the Various Diagnostic Groups on Demographic Variables.

Variable	ES (n=41)	PNES (n=43)	IS (n=44)	P
Age in years (SD)	36.9 (14.4)	40.6 (10.2)	37.3 (12.0)	0.311
Education in years (SD)	12.6 (2.3)	12.4 (2.6)	13.0 (2.3)	0.543
Gender (% females)	44.3[b]	79.1[a]	68.2[a]	**0.006**
Race (% Caucasian)	94.3	88.4	84.1	0.105
Handedness (% right)	85.7	88.4	90.9	0.731

Note: ES = Epileptic Seizures; PNES = Psychogenic Non-Epileptic Seizures; IS = Indeterminate Spells. Matching superscripts indicate that samples did not differ significantly on post-hoc comparison. Analysis of variance (ANOVA) was used to compare groups on age and education, while Kruskal-Wallis procedures were used to compare groups on gender, race, and handedness.

ES, PNES, and IS groups clearly failed the WMT at different rates (χ^2 (2, $\underline{N}$ = 128) = 12.96, $\underline{p}$ < .01). In fact, patients who failed the WMT were nearly five times more likely to be diagnosed with PNES or IS rather than ES (OR = 4.97, $\underline{p}$ < .001).

TTable 2. Frequency of Symptom Validity Test Failure Within the Various Diagnostic Groups.

	ES	PNES	IS	χ^2	P
# of pts failing WMT	7/41 (17.1%)[a]	22/43 (51.2%)[b]	22/44 (50.0%)[b]	12.96	**<.01**
# of pts failing WMT, excluding false positives	3/37 (8.1%)[c]	22/43 (51.2%)[d]	21/43 (48.8%)[d]	19.28	**<.001**

Note: WMT = Word Memory Test; ES = Epileptic Seizures; PNES = Psychogenic Non Epileptic Seizures; IS = Indeterminate Spells. Matching superscripts indicate that samples did not differ significantly on post-hoc comparison.

Of note, patients from the PNES and IS groups obtained by far the lowest scores on the effort-sensitive measures of the WMT. Six PNES and two IS patients actually scored below chance on the WMT (i.e., 13.9% of the PNES group and 4.5% of the IS group). In contrast, only one patient with ES performed within the range expected by chance alone, and this individual was one of the false positive cases who had never lived independently after sustaining a severe head injury.

Consistent with expectations, each group scored in the abnormal range on approximately half of the neurocognitive measures included in the Dodrill Discrimination Index (DDI), and the groups did not differ significantly. However, this changes if one views performance in light of the WMT results. Stratifying neurocognitive performance by WMT performance (pass/fail) reveals significant group differences in the DDI (F (5, 114) = 9.16, $p < .001$).

Table 3. Performance on the Dodrill Discrimination Index (DDI) by Diagnostic Group and WMT Performance.

	ES (n=40)		PNES (n=37)		IS (n=43)	
DDI (SD)	51.4 (24.9)		52.6 (27.7)		49.8 (28.0)	
WMT	Pass (n=34)	Fail (n=6)	Pass (n=19)	Fail (n=18)	Pass (n=22)	Fail (n=21)
DDI (SD)	49.4 (24.2)[a]	62.7 (28.1)[a,c]	33.1 (17.6)[b]	73.3 (20.6)[c]	36.1 (19.5)[b]	64.0 (28.8)[c]

Note: DDI = Dodrill Discrimination Index; WMT = Word Memory Test; ES = Epileptic Seizures; PNES = Psychogenic Non-Epileptic Seizures; IS = Indeterminate Spells. One patient from the ES group, six patients from the PNES group, and one from the IS group did not complete enough of the neurocognitive battery to produce a reliable DDI score. Matching superscripts indicate that samples did not differ significantly on post-hoc comparison.

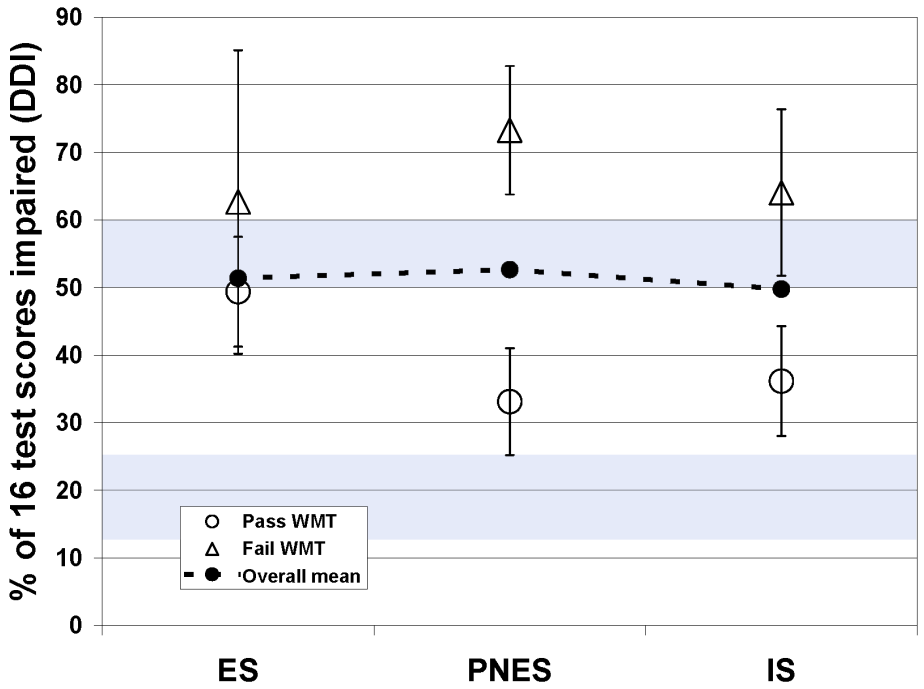

Figure 2. Mean Dodrill Discrimination Index (DDI) and 95% confidence intervals for patients with epileptic seizures (ES), psychogenic nonepileptic seizures (PNES), and seizures of indeterminate origin (IS), stratified by performance on the effort-sensitive measures of the Word Memory Test (WMT). A higher DDI value indicates more impairment. The dotted line depicts the mean level of performance for each group before stratifying for WMT performance. The uppermost shaded area is the 95% confidence interval of the DDI performance of 100 epilepsy patients (4), whereas the lower shaded area is the 95% confidence interval around the DDI performance of 50 normal control subjects (15).

The majority of the ES-Fail group was made up of patients who were not able to live in an independent fashion. This means that the PNES-Fail group, all of whom reported that they continued to live independently, scored in the same range as the most impaired patients with verified epilepsy.

Table 4. Odds Ratios for Patients who fail the WMT scoring in the Abnormal Range on 50% or More of the Neuropsychological Battery for Epilepsy by Diagnostic Group

Group	Odds Ratio	95% confidence interval	p value
ES	2.25	0.27 – 27.48	0.66
PNES	63.75	5.63 – 2833.38	< 0.0001
IS	5.36	1.23 – 24.30	0.01

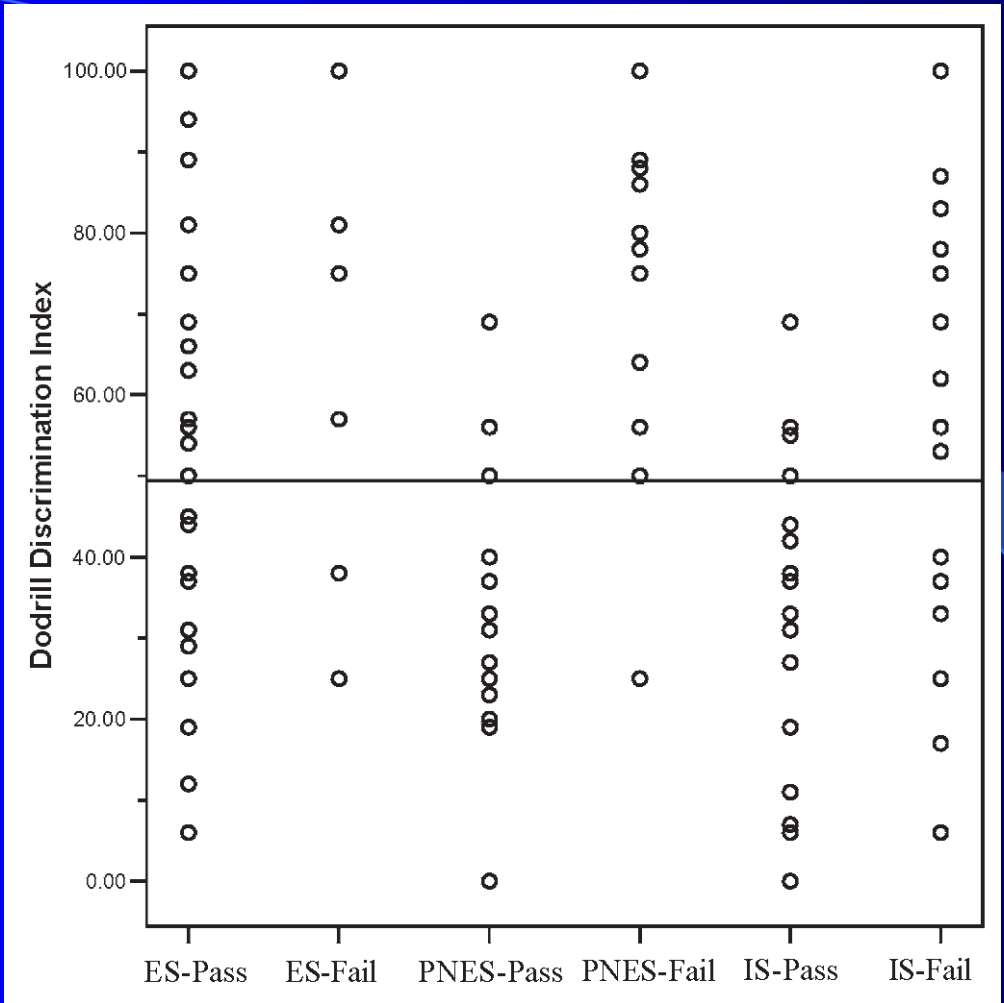

Figure 1. Scatterplot of DDI performance according to diagnostic group and WMT performance. DDI = Dodrill Discrimination Index; WMT = Word Memory Test; ES = Epileptic Seizures; PNES = Psychogenic Non-Epileptic Seizures; IS = Indeterminate Spells; Pass = scored in the valid range on WMT effort-sensitive tests; Fail = scored in the invalid range on WMT effort-sensitive tests. The horizontal reference line marks the mean performance of the ES group who passed the WMT.

Table 5. Comparison of Diagnostic Groups on Medical and Psychiatric Variables.

Variable	ES (n=41)	PNES (n=43)	IS (n=44)	P
Age at Spell Onset (in years) (SD)	19.0 (15.8)[a]	28.0 (13.4)[b]	25.4 (16.2)[b]	**< .03**
Frequency of Spells (per month) (SD)	17.0 (18.0)	21.6 (28.0)	31.0 (34.6)	.07
Number of Current AEDs (SD)	1.9 (0.8) [a]	1.2 (1.0) [b]	1.3 (1.0) [b]	**< .01**
Duration of Spells (years since onset) (SD)	17.5 (14.4)	12.8 (13.8)	11.9 (14.0)	.16
Focal Neurologic Exam	5/41 (12%)	3/43 (7%)	4/43 (9%)	.28

Table 5. Comparison of Diagnostic Groups on Medical and Psychiatric Variables.

Variable	ES (n=41)	PNES (n=43)	IS (n=44)	P
MRI abnormality	24/37 (65%)[a]	8/30 (27%)[b]	10/34 (29%)[b]	**<.001**
Baseline EEG findings	23/41 (56%)[a]	6/43 (14%)[b]	4/43 (9%)[b]	**<.01**
History of Neurologic Insult (self-report)	26/41 (63%)	29/43 (67%)	29/43 (67%)	.90
Neurologic History (consensus)	19/40 (47%)	18/43 (42%)	21/43 (49%)	.79
Psychiatric History	28/41 (69%)	35/43 (81%)	31/43 (72%)	.37

Table 5. Comparison of Diagnostic Groups on Medical and Psychiatric Variables.

Variable	ES (n=41)	PNES (n=43)	IS (n=44)	P
History of Closed Head Injury (self-report)	17/41 (41%)	27/43 (63%)	24/43 (56%)	.28
History of Sexual Abuse	8/41 (19%)	18/43 (42%)	16/43 (37%)	.15
History of Physical Abuse	10/41 (24%)[a]	24/43 (56%)[b]	15/42 (36%)[a]	**.05**
History of Emotional Abuse	13/41 (32%)	23/43 (53%)	19/41 (46%)	.28
History of Suicide Attempts	16/41 (39%)	15/38 (38%)	14/43 (33%)	.79

Table 5. Comparison of Diagnostic Groups on Medical and Psychiatric Variables.

Variable	ES (n=41)	PNES (n=43)	IS (n=44)	P
History of Suicide Attempts	16/41 (39%)	15/38 (38%)	14/43 (33%)	.79
History of Fibromyalgia	0/41 (0%)[a]	9/42 (21%)[b]	6/42 (14%)[b]	**<.01**
History of Chronic Pain	1/40 (2%)[a]	14/42 (33%)[b]	9/42 (21%)[b]	**<.01**

OOur data are consistent with our primary hypotheses. Specifically:

(1) Significantly more patients with PNES than those with ES failed symptom validity testing, suggesting that statements about neurocognitive performance of patients with PNES are confounded by invalid data,

22) Patients with PNES who performed within expected limits on the WMT significantly outperformed patients with ES on a well-validated neurocognitive battery previously shown to be insensitive to cognitive differences between the two groups, and

(3) Patients with PNES reported histories of fibromyalgia and chronic pain disorder more frequently than those with ES, and equivalent rates of neurologic disease and injury, although they showed less objective evidence of actual brain dysfunction.

Remaining Challenges in the Use of SVT in Neuropsychological Assessment:

11) Determining baserates of invalid effort for various diagnostic groups (e.g., mild TBI, epilepsy, stroke, psychiatric popuatlions).

(2) Identifying "false positives" on SVT measures.

X3) Exploring the impact of "mediating" factors (e.g., impact of psychiatric factors such as depression or anxiety, sleep deprivation/fatigue, pain, psychotropic medications, acute seizures).

(4) Exploring the impact of procedural factors in the use of SVTs (e.g., order of administration, whether or not to provide warnings, decision to discontinue testing).

Symptom Validity Testing (SVT): General Update on Current Practice Parameters and Specific Application in a Clinical Population

Daniel L. Drane, Ph.D., ABPP (CN)
Assistant Professor of Neurology
Director, Neuropsychology Laboratory
UW Regional Epilepsy Center