EVALUATION OF THE NEUROBEHAVIORAL SCREENING TOOL IN CHILDREN WITH FETAL ALCOHOL SPECTRUM DISORDERS (FASD)

Michael-Anne LaFrance[1], Kaitlyn McLachlan[1], Kelly Nash[2,3], Gail Andrew[1,4], Christine Loock[5,6] Tim F Oberlander[5,6], Gideon Koren[2,3], Carmen Rasmussen[1,4]

[1]Department of Pediatrics, University of Alberta; [2]Motherisk Program, The Hospital for Sick Children, Toronto; [3]Department of Pediatrics, University of Toronto; [4]Glenrose Rehabilitation Hospital, Edmonton; [5]Department of Pediatrics, University of British Columbia; [6]Child & Family Research Institute, University of British Columbia

ABSTRACT

Background
There is a growing need for validated tools to screen children at risk of fetal alcohol spectrum disorders (FASD). The Neurobehavioral Screening Tool (NST) is one of several promising screening measures for FASD, though further evidence is needed to establish the tool's psychometric utility.

Objective
To assess the predictive accuracy of the NST among children with an FASD diagnosis, with prenatal alcohol exposure (PAE) but no FASD diagnosis, and typically developing controls.

Method
The NST was completed by caregivers of children ages 6 to 17, including 48 with FASD, 22 with PAE, and 32 typically developing non-exposed controls. Predictive accuracy coefficients were calculated using Nash et al. (2006) criteria, and compared against controls. An alternative scoring scheme was also investigated to determine optimum referral thresholds using item-level total scores.

Results
The NST yielded 62.5% sensitivity for participants with FASD and 50% for PAE. Specificity values were 100% with no typically developing control scoring positive. Within the FASD group there was a trend for higher sensitivity among adolescents aged 12 to17 (70.8%) compared with children aged 6 to 11 years (54.2%), $p = 0.23$.

Conclusion
The findings support a growing body of literature evidencing psychometric promise for the clinical utility of the NST as an FASD screening tool, though further research on possible age-effects is warranted. The availability of a validated clinical screening tool for FASD, such as the NST, would aid in accurately screening a large number of children and lead to a timelier diagnostic referral.

Key Words: *Neurobehavioral screening tool, fetal alcohol spectrum disorder, prenatal alcohol exposure, screening*

Prenatal exposure to alcohol is a significant health concern often requiring extensive social and medical services for affected individuals and their families. Fetal alcohol spectrum disorders (FASD) is an umbrella term encompassing specific diagnoses that describe individuals who have physical, cognitive, and behavioral disabilities due to prenatal exposure to alcohol.[1] Prevalence estimates for FASD range between 1 to 5%.[1,2] In addition to significant neuropsychological deficits, individuals with FASD often have high rates of behavior problems, mental health issues,

and poor adaptive outcomes.[3] Children and youth with FASD may also experience adverse outcomes including school disruption, contact with the mental health and justice systems, and substance abuse problems.[4-8]

Until recently, there has been no screening tool for FASD due to a lack of identifiable biological markers or unique profile of cognitive and behavioral effects associated with FASD.[9] The Public Health Agency of Canada (PHAC) has published a National Screening Tool Kit for FASD, comprising five promising screening tools: The Neurobehavioral Screening Tool (NST), Meconium Fatty Acid Ethyl Esters Testing, Maternal Drinking Guide, Medicine Wheel Student Index/Developmental History, and the FASD Screening and Referral Form for Youth Probation Officers. [9-11]

The NST is a caregiver report questionnaire with ten items, which were extracted from the Child Behavior Checklist[12], representing common areas of behavioral concerns in children with FASD[11] The developers of the NST have published two studies assessing its psychometric utility, demonstrating the tool is able to statistically differentiate children with FASD from typically developing children and those with ADHD but no alcohol exposure. Nash and colleagues[10] identified unique patterns of items among children with FASD (6-16 years) that could differentiate them from unexposed children with ADHD including, lack of guilt after misbehaving, cruelty, and a tendency to act young for their age. These findings were replicated in a second study conducted by Nash, Koren and Rovet[13], where they found one item had important utility (acting younger than his or her age) in differentiating children with FASD from those with Oppositional Defiant Disorder (ODD) or Conduct Disorder (CD). Early NST research was limited by extracting the items from the administered CBCL protocols[10] and the use of retrospective chart reviews to extract parent responses from completed CBCL protocols.[13] However, the NST is now publicly available.

We sought to evaluate the screening properties of the NST in a cohort of children and adolescents previously diagnosed with an FASD, children with PAE who were assessed for FASD but not found to meet the threshold for diagnostic criteria, and typically developing children without prenatal alcohol-exposure. Establishing the NST's ability to discriminate among children with PAE who did not meet criteria for a diagnosis of FASD is important to further establish the specificity and discriminant validity of the NST among the full range of children with prenatal alcohol exposure. We also examined whether possible gender and age differences existed with respect to the tool's psychometric parameters (e.g., sensitivity, specificity). Results of this study also provide the first data on the direct implementation of the NST as a separate tool to caregivers and may lead to further validation of the NST as a neurobehavioral screening tool for FASD.

MATERIALS AND METHODS

Participants

Participants included 102 children and adolescents ranging in age from 6 to 17 *(M* = 12.00 *SD* = 2.92). In total, 48 had an FASD diagnosis, 22 had confirmed PAE and were assessed for FASD but did not meet diagnostic thresholds at the time of their assessments, and 32 were typically developing controls (see Table 1). Both clinical and non-clinical samples were recruited from the Edmonton and Vancouver areas. Participants in Edmonton were drawn from the FASD Clinical Services program at the Glenrose Rehabilitation Hospital in Edmonton. Children from the Vancouver cohort comprised a subset of participants recruited for the NeuroDevNet FASD demonstration study[14] and were assessed at a variety of FASD clinics around the lower mainland area. Typically developing controls were recruited from schools, community centres, and other local health centres in the same geographic regions as the PAE and FASD groups.

TABLE 1 Participant Characteristics

	FASD *n* = 48	PAE *n* = 22	Control *n* =32	*p*-value
Sex (% female)	56.3%	36.3%	68.8%	0.062 (ns)[a]
Mean age (range)	12.17 (6.83-17.92)	11.50 (7.08-17.67)	12.00 (6.92-17.17)	0.689 (ns)[b]
Current living arrangement				
Biological family	20.8%	13.6%	96.9%	0.000 [a]
Adopted	39.6%	31.8%	0%	
Foster care	22.9%	27.3%	3.1%	
Kinship	16.7%	27.3%	0%	
Mean number of living situations (range)	3.6 (1-9)	3.4 (1-10)	1.16 (1-3)	0.000 [b]
Ethnicity				0.000[a]
Caucasian	27.1%	9.1%	65.6%	
Aboriginal	62.5%	72.1%	0%	
Other/Mixed	10.4%	18.2%	34.4%	
Mean SES[a] (SD)	34.5 (14.3)	36.2 (13.3)	44.3 (10.9)	0.005[b]
One or more comorbid diagnoses	81.3%	61.8%	6.3%	
Most common diagnoses				
ADHD	60.4%	40.9%	-	
Language disorder	20.8%	9.1%	-	
Anxiety or Depressive Disorder	20.8%	18.2%	6.3	
Learning Disorder	6.3%	18.2%	-	
Reactive Attachment Disorder	8.3%	13.6%	-	
Current medication[d] use (%)	54.2%	36.4%	-	

Note: Social economic status (SES) was calculated using Hollingshead's Four-Factor Index of Social Status (Hollingshead, 1975). [a]Analyzed by chi-square analysis; [b]analyzed by ANOVA. ADHD = Attention Deficit Hyperactivity Disorder. c SES = Socioeconomic status [d] Medication use assessed included only psychiatric and seizure based pharmaceuticals.

Across study sites, children in both the FASD and PAE groups were assessed by clinical teams who follow the Washington Diagnostic and Prevention Network's (FAS DPN) 4-Digit Diagnostic Code[15], and the Canadian Guidelines for FASD Diagnosis.[1] The 4-Digit Diagnostic Code uses a 4-point Likert scale to assign severity ranking across four indicators, including growth deficiency, facial phenotype, central nervous system (CNS) dysfunction and alcohol use.[15] Both prenatal (e.g., poor prenatal care, concurrent substance exposure) and postnatal (e.g., neglect, caregiver disruption) events and exposures are also ranked. All children in the FASD and PAE groups had a reliable source confirm prenatal exposure to alcohol (all ranked 3 or 4), and assessments were typically conducted by a multidisciplinary team.

Measures

The Neurobehavioral Screening Tool

The NST[10] is a parent/caregiver self-report measure assessing ten common behavioral concerns expressed by caregivers of children with FASD. The tool is intended to screen children ages 6 through 18 at risk of neurobehavioral problems following prenatal exposure to alcohol. The NST consists of ten items drawn from the Child Behavior Checklist (CBCL)[12] a commonly used clinical measure of behavior and social

competence in children ages 6 to 18 years following research demonstrating they were highly discriminant of children with an FASD.[10,16]

The NST is meant to be completed within the context of a clinical interview and requires the respondent to have familiarity with the child's behavior over the past six-months. Caregivers provide a "yes" or "no" response to each item, and the instrument typically requires 5-minutes to complete. The NST was scored using a two-step referral algorithm[10], where step 1 identifies behavior suggestive of FASD ('yes' to at least 6 items of 1-7 or 'yes' to at least 3 of items 1, 3, 4, and 5) and step 2 differentiates FASD from ADHD ('yes' to 2 items 1, 5, 8 or 3 items 1, 5, 8, 9, 10). Children must meet criteria in both steps to receive a positive screen. Nash and colleagues have demonstrated high sensitivity (86% through 98%) and variable specificity (42% through 82%) using these criteria in two samples of children and adolescents against typically developing controls.[10,13]

Procedure

This study involved retrospective administration of the NST by caregivers. Caregivers of those in the FASD and PAE groups completed the NST along with other rating scales as part of a larger study and it was not conducted within a clinical interview. Caregivers of healthy controls in Edmonton did not attend study sessions and the NST was mailed out for home completion with instructions and contact information for the research team. Caregivers also provided demographic and clinical information about their child either during a semi-structured interview or by home questionnaire. All study procedures were approved by the University Research Ethics Review Boards at the University of Alberta and the University of British Columbia, and adhered to governing ethical guidelines.

Data Analysis

To assess the possibility that SES had a differentially confounding effect on NST screening outcome by site, a moderated hierarchical logistic regression was conducted by entering SES and site in the first block, and the interaction between the SES and site in the second block.[17] The interaction was not significant (p = .13), indicating SES was not differentially related to NST screening outcome by site. Further, point-biserial correlations between SES and NST screening outcomes were not significant (p = .23).

Therefore, it was possible to collapse across samples in the following analyses. NST performance indicators were calculated using MedCalc,[18] an online tool developed to assist in the evaluation of diagnostic and screening measures. *Sensitivity* refers to the probability that a test result will be positive when the disease is present, whereas *specificity* refers to the probability that a test result will be negative when the disease is not present. In general, the higher the sensitivity and specificity values, the better, recognizing that higher values on one side or the other may result in an accuracy trade off. *Positive predictive value* (PPV) measures the probability that a disease is present when the test is positive, while *negative predictive value* (NPV) reflects the probability that disease is absence when the test is negative. PPV and NPV are critical indicators of screening test performance because they allow us to predict how accurately a positive test result will lead to an FASD assessment and diagnosis in the general population in which diagnostic history is unknown. Contrasts among the FASD, PAE, and control groups on individual NST items were conducted using Chi-square analyses. One-way ANOVA analyses were used to compare the three participant groups on demographic characteristics

RESULTS

Participant Characteristics

Among the FASD group, 23 (47.9%) were diagnosed with static encephalopathy, 17 (35.4%) with neurobehavioral disorder, 5 (10.4%) with partial fetal alcohol syndrome, and 3 (6.3%) with fetal alcohol syndrome according to the Washington FAS DPN four-digit code.[15] Participants from both the FASD and PAE groups were predominantly under the care of adoptive, foster or kinship caregivers, whereas the majority of the typically developing controls lived with their biological parents. Children in the FASD and PAE groups also had a higher average number of lifetime living situations and lower socioeconomic

status (SES) relative to controls. These differences are all consistent with research involving children with FASD, owing largely to the social determinants of health that typically underlie alcohol consumption during pregnancy.[19,20] Children in both the FASD and PAE groups presented with a wide range of comorbid mental health conditions (Table 1) with the most common being ADHD, which is typical of this population.[5,6,21,22]

Demographic characteristics of participants between sites did not differ significantly, with the exception of SES; participants with FASD from Edmonton had significantly higher SES scores (M = 37.49, SD = 13.53) compared to those from Vancouver (M = 24.45, SD = 12.67), t (46) = 2.84, p = .007. Participants were eligible to participate in the control group if a caregiver was able to confirm the child was not exposed to alcohol prenatally and did not have a significant neurological or cognitive diagnosis (epilepsy, brain injury)).

NST Screening Outcomes

Using the scoring algorithm published by Nash et al.,[10] nearly two-thirds of children with an FASD (n = 30) screened positive on the NST, compared to none of the typically developing controls, resulting in an overall sensitivity rate of 62.5%, specificity of 100.0%, 100.0% PPV, and 64.0% NPV. Half of participants in the PAE group (50.0%) screened positive on the NST, producing similar performance coefficients, including 50.0% sensitivity, 100% specificity, and 100% PPV, and 74.4% NPV (sensitivity values did not differ significantly between the PAE and FASD groups, p = .39).

Descriptive data detailing item level endorsement rates for each item are presented in Table 2. Significantly higher endorsement rates were seen in the FASD group compared to typically developing controls on all ten NST items. Children with PAE also showed significantly higher endorsement rates than controls on all but two items (#8: cruelty, bullying and meanness to others, and, #10: theft of items from outside the home). Overall, several items were endorsed at much lower frequencies among the three groups, including items eight (bullying, cruelty, meanness to others), nine (steals items from the home), and ten (steals items from outside the home).

TABLE 2 Endorsement rates for FASD, PAE, and control groups on the 10 NST items

NST Item		Percentages			*p*-value		
		FASD	PAE	Control	FASD vs Controls	FASD vs PAE	PAE vs Controls
1	Has your child been seen or accused of or thought to have acted too young for his or her age?	72.9	72.7	3.1	0.000*	0.987	0.000*
2	Has your child been seen or accused of or is thought to be disobedient at home?	81.3	81.8	6.3	0.000*	0.955	0.000*
3	Has your child been seen or accused of or is thought to lie or cheat?	72.9	77.3	18.8	0.000*	0.699	0.000*
4	Has your child been seen or accused of or is thought to lack guilt after misbehaving?	70.8	36.4	0	0.000*	0.006*	0.000*
5	Has your child been or accused of or thought to have difficulty concentrating, and can't pay attention for long?	91.7	86.4	12.5	0.000*	0.492	0.000*
6	Has your child been seen or accused of or is thought to act impulsively and without thinking?	91.7	90.9	18.8	0.000*	0.916	0.000*
7	Has your child been seen or accused of or is thought to have difficulty sitting still is restless or hyperactive?	85.4	63.6	6.3	0.000*	0.039	0.000*
8	Has your child been seen or accused of or is thought to display acts of cruelty, bullying or meanness to others?	47.9	36.4	12.5	0.001*	0.366	0.038
9	Has your child been seen or accused of or is thought to steal items from home?	41.7	27.3	3.1	0.000*	0.247	0.009*
10	Has your child been seen or accused of or is thought to steal items outside of the home?	39.6	13.6	0	0.000*	0.030	0.032

Age and Gender Differences on the NST

In order to assess possible age differences using the NST, participants were divided into two age groups: children ages 6 through 11 (M = 9.22, SD = 1.54, n = 54, 52.9%) and adolescents ages 12 through 17 (M = 14.17, SD = 1.64, n = 48, 47.1%). There was a non-significant trend toward age differences in both the FASD and PAE groups (vs. controls). Specifically, there were more positive screens and higher sensitivity among adolescents in the FASD group (70.8%) compared to children (54.2%), χ^2 (1, N = 48) = 1.42, p =0.23). The same pattern was evident among children in the PAE group, with a higher positive screening rate and sensitivity among adolescents (71.4%) compared to children (40.0%), χ^2 (1, N = 22) = 1.89, p = 0.17 (Table 3).

TABLE 3 NST predictive accuracy among FASD and PAE participants vs. controls

Group	Sensitivity	Specificity	PPV	NPV
FASD	62.50	100.00	100.00	64.00
Children	54.17	100.00	100.00	57.69
Boys	60.00	100.00	100.00	84.62
Girls	50.00	100.00	100.00	75.86
Adolescents	70.83	100.00	100.00	70.83
Boys	81.82	100.00	100.00	91.67
Girls	61.54	100.00	100.00	81.48
PAE	50.00	100.00	100.00	74.40
Children	40.00	100.00	100.00	62.50
Boys	33.33	100.00	100.00	78.57
Girls	50.00	100.00	100.00	88.00
Adolescents	71.40	100.00	100.00	89.47
Boys	80.00	100.00	100.00	95.65
Girls	50.00	100.00	100.00	95.65

***Note*.** All groups are referenced against respective control subjects. Age-specific comparisons are referenced against comparably-aged controls (e.g., FASD young are compared with young controls). PPV = Positive Predictive Value. NPV = Negative Predictive Value.

To further investigate possible reasons underlying age-related differences in sensitivity, exploratory item-level analyses were also conducted. First, examining the entire sample, we found that two items (9, theft inside the home, and 10, theft outside the home) had roughly double the endorsement rate in older adolescents relative to younger children, but these trends did not reach statistical significance (p = .054, and p = .079, respectively). Within the FASD group, two items (9 and 10) were endorsed at substantially lower rates among younger children compared to adolescents, as well as items four and nine among younger children with PAE (see Table 4). Across groups the trend saw higher endorsement rates for adolescents compared to younger children on almost all of the 10 items, none of them however reached significance at p-value <0.01.

TABLE 4 Percentage endorsement rates for children and adolescents in FASD, PAE, and control groups

NST Items		FASD		PAE		Controls	
		6-11 (*n* = 24)	12-17 (*n* = 24)	6-11 (*n* = 15)	12-17 (*n* = 7)	6-11 (*n* = 15)	12-17 (*n* = 17)
1	Acted too young for his or her age?	62.5	83.3	80.0	57.1	0	5.9
2	Disobedient at home?	79.2	83.3	80.0	85.7	6.7	0
3	Lie or cheat?	75.0	70.8	66.7	100	20.0	17.6
4	Lack guilt after misbehaving?	66.7	75.0	20.0	71.4	0	0
5	Difficulty concentrating, and can't pay attention for long?	87.5	95.8	86.7	85.1	6.7	17.6
6	Act impulsively and without thinking?	87.5	95.8	86.7	100	13.3	23.5
7	Difficulty sitting still is restless or hyperactive?	79.2	91.7	60.0	71.4	0	11.8
8	Display acts of cruelty, bullying or meanness to others?	54.2	41.7	33.3	42.9	6.7	17.6
9	Steal items from home?	25.0	58.3	20.0	42.9	6.7	0
10	Steal items outside of the home?	25.0	54.2	13.3	14.3	0	0

Chi-square analyses indicated no significant differences. All p = >.01.

Trends toward gender differences were also evident but did not reach significance. The NST showed somewhat higher sensitivity among boys in the FASD group (71.4%) compared to girls (55.6%). Alternatively, rates were equal between boys (50.0%) and girls (50.0%) in the PAE group. When we looked at age by gender interactions, older boys in both the FASD (81.8%) and PAE (80.0%) groups showed higher sensitivity on the NST compared to younger boys (60.0% for FASD and 33.3% for PAE), whereas rates were approximately even among younger and older girls in both groups. An item-level analysis revealed few clear trends in gender differences by item. Within the FASD group, overall, boys endorsed higher rates than girls on items 1 (acts too young), 2 (disobedient at home), 4 (guilt after misbehaving), 5 (can't pay attention), 6 (impulsive), 8 (bullying, meanness to others), 9 (steals from home) and 10 (steals outside home), though only differences on item four reached significance (85.7% vs. 59.3%, χ^2 (1, N = 48) = 4.00, p = .045. Girls had higher endorsement rates than boys on items 3 (lying or cheating), and 7 (restless, hyperactive), though differences were not significant. This trend was reversed in the PAE group, with girls demonstrating higher endorsement rates on items 2 (disobedient at home), 3 (lying or cheating), 5 (can't pay attention), 6 (impulsive), 7 (restless, hyperactivity), 9 (steal from home), and 10 (steal outside home). Across all items, boys in the control group had higher item level endorsement rates than girls, suggesting possible gender differences in patterns of NST item endorsement across groups.

Exploratory Cumulative Scoring Option

Differences in item level endorsement appeared to influence screening accuracy using the Nash et al.[10] scoring algorithm, thus, we examined an alternative method for scoring the NST. Specifically, we evaluated optimum screening cut-points at each level of cumulative item

endorsement (irrespective of positioning on the NST, see Figure 1a), allowing us to examine optimum sensitivity and specificity values among both younger children and adolescents in this sample (Figure 1b, 1c). In the present sample, the "optimal" screening threshold (determined by assessing the best balance between achieving high sensitivity without greatly sacrificing specificity) appeared to be any four or more items endorsed positively on the NST, whereas moving to five items reduced sensitivity among the FASD group without changing sensitivity among the PAE group or specificity. Similarly, selecting only three items produced a considerable loss in specificity. Among participants with FASD, the cut point of four items produced a sensitivity of 89.6%, and 90.6% specificity (Figure 1a). Although age related differences were again present (Figure 1b and 1c), among the FASD group, screening accuracy (using a cut point of four items endorsed) was higher among both younger children (sensitivity = 83.3%, specificity = 100.0%) and adolescents (sensitivity = 95.8%, specificity = 82.4%) compared with the traditional scoring algorithm, although sensitivity was poorer for those 12 years and older.

FIG. 1A Percentage of children screening positive on the NST based on number of cumulative items endorsed

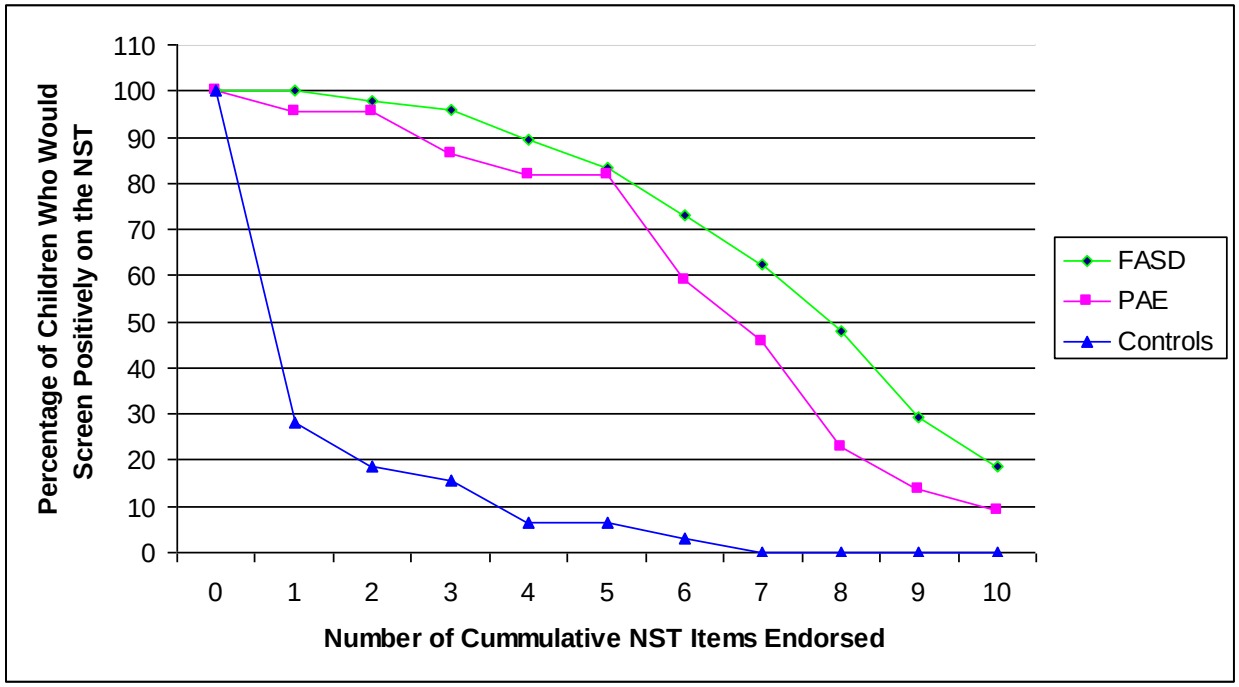


FIG. 1B Percentage of children under 12 years screening positive on the NST based on number of cumulative items endorsed

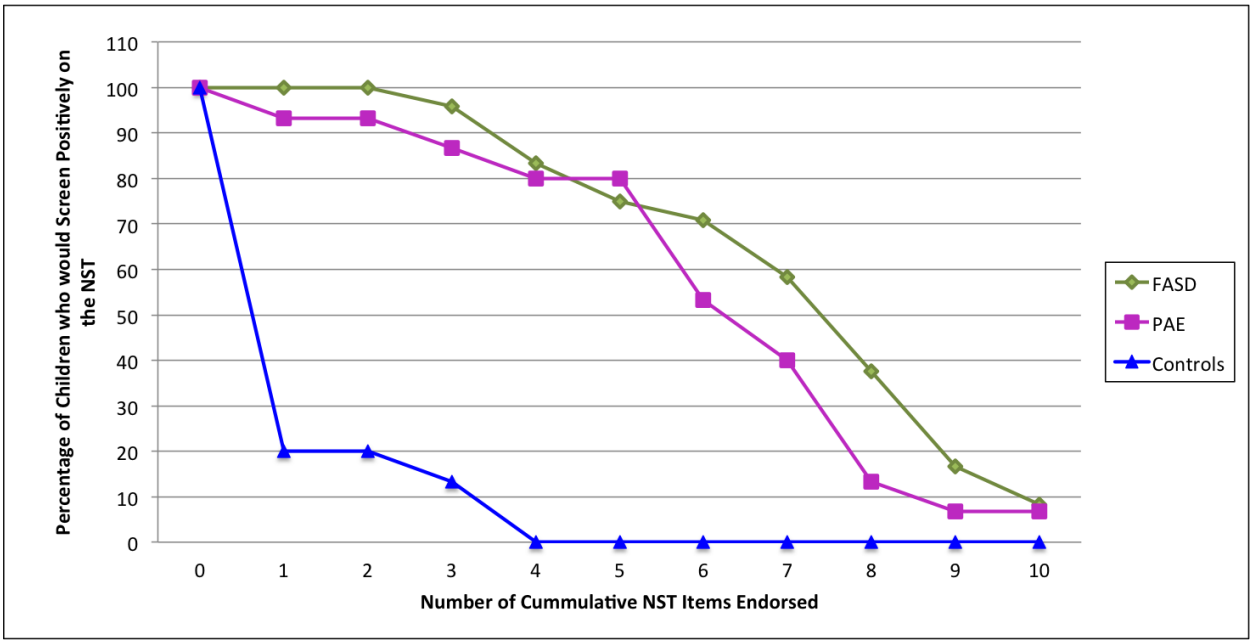
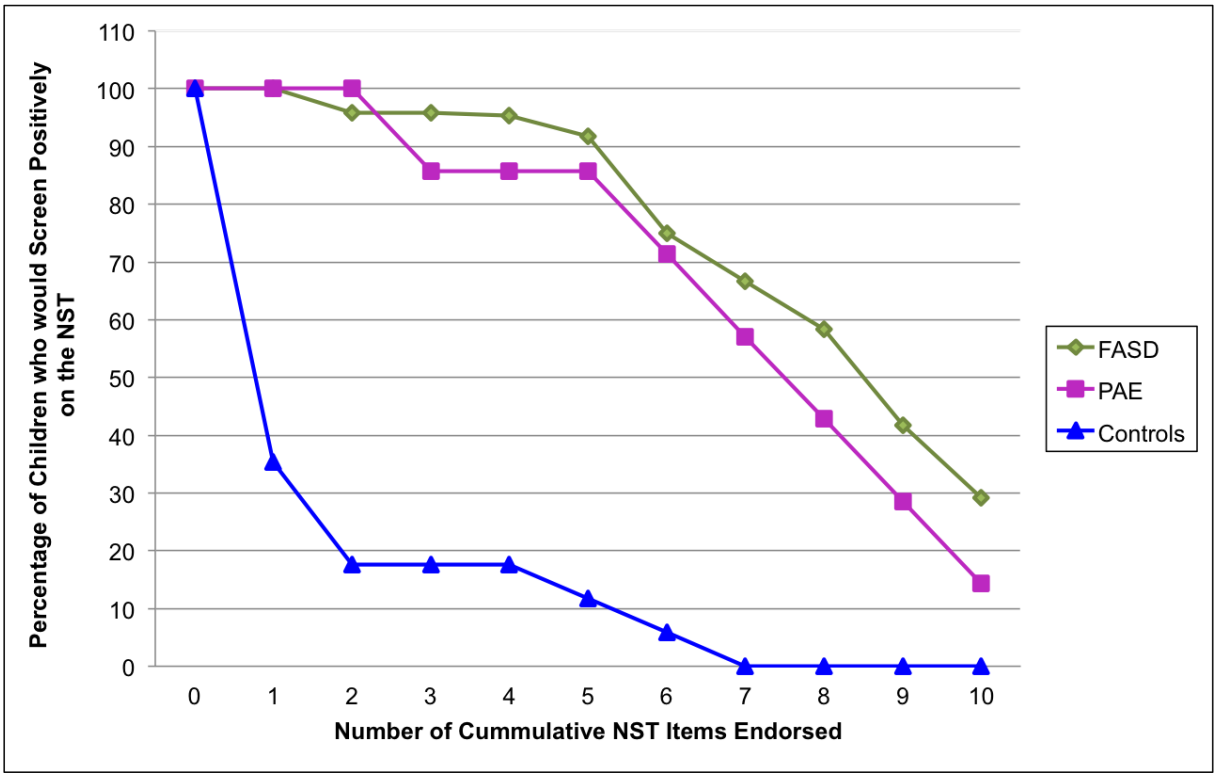


FIG. 1C Percentage of children 12 years and older screening positive on the NST based on number of cumulative items endorsed



DISCUSSION

The goal of this study was to evaluate the screening performance of the NST among children with an FASD diagnosis, with PAE but no FASD diagnosis, and typically developing controls. To our knowledge, this study is the first to implement the NST, in its published form, with parents completing the measure as a stand-alone tool, having an alcohol-exposed non-diagnosed group, and examining age and gender on the NST. In spite of high demand, access to diagnostic services for children and adolescents with PAE remains highly limited in Canada[23] and thus it is critical that appropriate screening measures be validated to assist in triaging children and youth who would most benefit from further assessment and intervention. Although early research has demonstrated promising findings with respect to the NST's performance as an FASD screening tool, it is critical to replicate and extend this work.

NST Sensitivity and Specificity

The present findings showed 100% specificity on the NST among typically developing controls, in keeping with earlier findings. However, we found somewhat lower sensitivity compared to the original studies by Nash and colleagues[10,13], with 62.5% of participants with FASD, and 50.0% of participants with PAE screening positive. Several reasons may explain these differences. First, tools optimized to achieve specific performance indicators in one sample, often yield different sensitivity in later samples[24], particularly if diagnostic practices differ slightly. This underscores the importance of cross validating screening tools for clinical use in multiple samples before drawing conclusions about the validity of an instrument such as the NST.

Children with FASD in both Nash cohorts included only those who presented with probable or definite CNS dysfunction. However, one-third of participants with FASD in the present study presented with CNS score of 2 (possible), because diagnosticians in our study locations follow the Washington FAS DPN diagnostic criteria[15], which permits diagnosis in certain cases where a CNS code of 2 is observed. In the current study, fewer than half (43.8%) of children in the FASD and PAE groups with a CNS score of 2 screened positive on the NST, whereas more than two thirds (71.9%) of those with CNS scores of 3 or 4 screened positive. Thus, it appears that the NST is more sensitive among individuals who present with higher levels of CNS dysfunction and should be explored in future research.

We administered the ten NST items in their published form instead of extracting items from completed CBCLs[10] or using CBCL data extracted from a chart review.[13] It is possible that caregivers had a differential response bias when completing the ten NST items together, rather than within the larger context of the CBCL. Also, the CBCL provides parents a three-point response scale ("Not True," "Somewhat True" and "Very True"). Though Nash and colleagues[10,13] included CBCL items that were endorsed as either "Somewhat True" or "Very True" in published studies, the final version of the NST offers caregivers only a binary "yes/no" response option. Caregivers may be been more prone to answer "no" if the child had only displayed the behavior a few times (compared to the "somewhat true" CBCL option). Further research using the NST in practice should examine possible differences in providing a 3-point versus dichotomous response scale, consistent with the CBCL.

Finally, in Nash et al.[10] caregivers completed the CBCL either in the context of a clinical assessment for FASD or during a follow-up assessment for children with ongoing behavioral problems following diagnosis. Alternatively, participants in our FASD and PAE groups were recruited at least one year after an FASD assessment, without having come forward to address concerns about behavioral problems. During this interval, it is likely that many families accessed recommended supports and services, further diminishing clinical need by the time participants were enrolled in the study. As such, behavioral problems may have diminished during the six-months prior to study enrolment, resulting in lower base rates of problem behaviors typically associated with FASD and PAE. Alternatively, because parents of all children in the FASD and PAE groups had been through the diagnostic process with their child, they likely held a higher level of knowledge about the typical behaviors

associated with FASD, as well as problematic behaviors within their own child based on clinical findings. This knowledge could have biased their ratings, relative to a general population of caregivers with limited knowledge about PAE-related behaviors and or/insight into their own child's behavioral profile. Prospective research administering the NST prior to an FASD referral or assessment would yield more unbiased findings and compliment the current results.

Age and Gender Differences

We found a trend toward better sensitivity better (approximately 15% better) among adolescents with FASD and PAE, relative to their younger counterparts. Age-related effects seemed to be weighted by endorsement rates on two items, including item nine ("steals items from home"), and ten ("steals items from outside the home"), which here higher among adolescents. The low base rate of endorsement on these items may suggest they are not predictive of the effects of prenatal exposure to alcohol in younger children. Research examining the onset of behaviors such as these and similar antisocial behaviors in the general population indicates variability in developmental onset.[25,26] Further, scoring the NST using a cumulative item approach resulted in higher sensitivity values, with a 40% increase among younger children. However this approach requires much further research to establish any possible sacrifices in the tool's ability to discriminate among other clinical disorders with overlapping behavioral profiles. Further research on the NST among early grade school children is important. The gender differences on the NST in the current study are in keeping with general findings from the developmental literatures showing higher rates of antisocial behaviors and conduct related behaviors reported among boys.[27-29] Though trends on gender differences did not reach significance, it will be important in future research to explore the possibility that the NST could under identify girls with PAE and/or FASD relative to boys given the behaviorally loaded and externalizing focus of NST items.

Children with PAE as a Screening Comparison Group

The addition of a PAE group further adds to our understanding of how the NST functions across the full spectrum of children with prenatal exposure to alcohol. These children with PAE may not be diagnosed with an FASD for a variety of reasons including: they did not demonstrate sufficient levels of neurobehavioral impairment or to meet diagnostic criteria for FASD, they were too young to be adequately assessed on all neurobehavioral domains, or perhaps had other significant life adversities or medical issues that precluded a diagnosis at the time of assessment. However, most were assessed by interdisciplinary diagnostic teams, had confirmed prenatal exposure to alcohol, and the majority presented with some degree of neurobehavioral difficulty (95.5% with a CNS code of two). Given the level of neurobehavioral challenges apparent in this group, it is possible that some children may actually fall on the FASD spectrum, later meeting criteria for FASD upon further assessment when neuropsychological and behavioral development is more solidified, and real world demands exceed capacities.

In general population screening settings the NST will need to capture children who have both PAE and neurobehavioral problems, however, many who are referred for full assessment will not qualify for a diagnosis on the FASD spectrum. There may not necessarily be a clear way to differentiate these children in advance. The fact that half of children in the PAE group did not screen positive for further assessment using the NST may be problematic. Further research is also needed to clarify the NST's ability to differentiate children with PAE from the profiles of other clinical populations such as ADHD and ODD/CD. It is also important to note that alcohol exposure was known for all children prior to their participation in this study, and all children with PAE had been referred for an alcohol-exposure related assessment. Results using the NST may be different in situations where alcohol exposure history is unknown and the tool is used among a broad sample of children who have not necessarily been identified as being at risk for problems related to PAE. It is therefore

difficult for us to conclude whether the tool is sensitive to PAE since all participants were previously referred for diagnosis and our sample did not include those where alcohol exposure unknown.

Limitations

This study involved retrospective administration of the NST in already assessed and diagnosed children, thereby limiting the extent to which findings can be said to establish the validity of the NST as a *screening tool*. That said our results contribute new and important information to the growing literature of studies assessing the sensitivity and specificity of the NST in a variety of populations and clinical settings. Our sample size was modest for this type of research, particularly our PAE group, which restricts the generalizability of our findings. Another important limitation that plagues much of the work with this population is that individuals with FASD are known to experience high rates of adverse prenatal and postnatal experiences, as well as significant caregiver disruption. These factors are also associated with the onset of problematic behavioral patterns in children who do not necessarily have PAE. Our "control" sample may be characterized as one free of many of the additional daily stressors experienced by children in our FASD and PAE groups, and it is not surprising that very few children in this group exhibited problematic behavioral patterns on the NST. In addition, limited sample size may have adversely affected the statistical power necessary to detect site by SES interactions in analyses assessing the impact of SES on NST results by group.

It is also important to recognize that irrespective of early data showing strong specificity of the NST among typically developing children relative to clinical groups without PAE, the NST may nevertheless incorrectly identify children with behavioral problems that are FASD-consistent, but who do not have prenatal exposure to alcohol. Although this may be seen as a limitation of the tool, children with serious behavioral problems would nevertheless benefit from clinical evaluation and supportive services to address these needs, and likely would not be specifically assessed for FASD in the absence of confirmed prenatal exposure to alcohol. Further research examining the NST among children with identified behavioral problems, as well as prenatal and postnatal stressors, but without prenatal exposure to alcohol is critical, along with replication of the current findings in a larger cohort of children varying in socioeconomic backgrounds.

CONCLUSION

The NST has promising potential as a tool for quickly and effectively screening a large number of children and adolescents. The availability of valid and reliable FASD screening tools will serve a critical role in both helping to triage overburdened assessment clinics that do not presently have access to such tools, and aid in the identification of children with PAE who may not otherwise come to the attention of clinicians. It is important to emphasize that the NST is intended for screening purposes *only* to identify those children and adolescents who may be in most need of a full clinical assessment and should not be used as a diagnostic tool.[13] The NST is still in the validation stage, and our findings lend support to the advisability of continued testing before full implementation of the tool, particularly among younger children.

Implementation of the NST and other FASD screening tools must also come with important consideration of ethical principles. Indeed, implementation of the NST on a large-scale, such as within the primary school system, may result in even higher numbers of children being referred, further burdening clinics with unmanageable waitlists and highlighting the lack of clinical support services for FASD. Consideration about the need for increased funding and supports for assessment and intervention services will form an important element of the conversation involved with policy makers in order to meet increasing demands use. It is also critical that caregivers be advised that a positive screening result on the NST is not necessarily indicative of a later FASD diagnosis, as is evidenced by the present findings. In spite of these concerns implementation of a validated

version of the NST could have important social impact, potentially leading to earlier identification and assessment thus providing more opportunities for remediation, which could lead to a reduction in adverse outcomes for affected children and families. Implementation of the NST may also aid in facilitating the determination of truer prevalence rates of FASD in Canada.

Thus, further validation efforts of the tool are encouraged, in combination with collaborative discussion with clinicians and policy makers responsible for allocating resources.

Sources of Support: For all reprint requests and other correspondences please contact Public Health Agency of Canada.

Corresponding Author: Carmen Rasmussen
carmen@ualberta.ca

REFERENCES

1. Chudley A, Conry J, Cook J, Loock C, Rosales T, LeBlanc N. Fetal alcohol spectrum disorder: Canadian guidelines for diagnosis. Canadian Medical Association Journal. 2005;172:S1-S21.
2. May P, Gossage J, Kalberg W, et al. Prevalence and epidemiologic characteristics of FASD from various research methods with an emphasis on recent in-school studies. Developmental Disabilities Research Reviews 2009;15:176-192.
3. Rasmussen C, Andrew G, Zwaigenbaum L, Tough S. Neurobehavioral outcomes of children with Fetal Alcohol Spectrum Disorder: A Canadian perspective. Pediatrics & Child Health 2008;13(3):185-191.
4. Fast D, Conry J, Loock C. Identifying fetal alcohol syndrome among youth in the criminal justice system. Journal of Developmental and Behavioral Pediatrics 1999;20:370-372.
5. O'Connor MJ, Shah B, Whaley S, Cronin P, Gunderson B, Graham J. Psychiatric illness in a clinical sample of children with prenatal alcohol exposure. American Journal of Drug and Alcohol Abuse Nov 2002;28(4):743-754.
6. Pei J, Denys K, Hughes J, Rasmussen C. Mental health issues in Fetal Alcohol Spectrum Disorder. Journal of Mental Health 2011;20(5):438-448.
7. Streissguth A, Bookstein F, Barr H, Sampson P, O'Malley K, Young J. Risk factors for adverse life outcomes in Fetal Alcohol Syndrome and Fetal Alcohol Effects. Journal of Developmental & Behavioral Pediatrics Aug 2004;25(4):228-238.
8. Yates W, Cadoret R, Troughton E, Stewart M, Guinta T. Effects of fetal alcohol exposure on adult symptoms of nicotine, alcohol, and drug dependence. Alcohol Clin Exp Res 1998;22(4):914-920.
9. Goh Y, Chudley A, Clarren S, et al. Development of Canadian screening tools for fetal alcohol spectrum disorder. The Canadian Journal of Clinical Pharmacology 2008;15(2):344-366.
10. Nash K, Rovet J, Greenbaum R, Fantus E, Nulman I, Koren G. Identifying the behavioural phenotype in fetal alcohol spectrum disorder: sensitivity, specificity and screening potential. Archives of Women's Mental Health 2006;9:181-186.
11. CAPHC. The Screening Tools. National Screening Tool Kit for Children and Youth Identified and Potentially Affected by FASD 2012; http://ken.caphc.org/xwiki/bin/view/FASDScreeningToolkit/National+Screening+Tool+Kit+for+Children+and+Youth+Identified+and+Potentially+Affected+by+FASD.
12. Achenbach T, Rescorla L. Manual for ASEBA school-age forms and profiles. Burlington, VT: University of Vermont, Research Centre for Children, Youth & Families; 2001.
13. Nash K, Koren G, Rovet J. A differential approach for examining the behavioral phenotype of fetal alcohol spectrum disorders. J Popul Ther Clin Parmacol 2011;18(3):e440-e453.
14. Reynolds JN, Weinberg J, Clarren S, et al. Fetal Alcohol Spectrum Disorders: Gene-environment interactions, predictive biomarkers, and the relationship between structural alterations in the brain and functional outcomes. Semin Pediatr Neurol 2011;18(1):49-55.
15. Astley S. Diagnostic Guide for Fetal Alcohol Spectrum Disorders:The 4-Digit Diagnostic Code. 3rd ed. Seattle, WA: University of Washington Publication Services; 2004.
16. Greenbaum R. Socioemotional functioning in children with alcohol related neurodevelopmental disorder (ARND): Profile on the Child Behavior Checklist (CBCL) [Unpublished Master's Thesis]. Toronto: Department of Human Development and Applied Psychology Ontario Institute for Studies in Eduation/University of Toronto; 2000.
17. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical

considerations. Journal of Personality and Social Psychology 1986;51:1173-1182.
18. MedCalc Software. MedCalc: Diagnostic test evaluation. 2013; http://www.medcalc.org/calc/diagnostic_test.php
19. Coggins TE, Timler GR, Olswang LB. A state of double jeopardy: impact of prenatal alcohol exposure and adverse environments on the social communicative abilities of school-age children with fetal alcohol spectrum disorder. Lang Speech Hear Serv Sch Apr 2007;38(2):117-127.
20. Olson HC, Oti R, Gelo J, Beck S. "Family Matters": Fetal alcohol spectrum disorders and the family. Developmental Disabilities Reviews 2009;15(3):235-249.
21. Rasmussen C, Benz J, Pei J, et al. The impact of an ADHD co-morbidity on the diagnosis of FASD. Can J Clin Pharmacol 2010 Winter;17(1):e165-76 Epub 2010 Apr 6. 2010.
22. Streissguth AP, Barr HM, Kogan J, Bookstein FL. Understanding the occurrence of secondary disabilities in clients with fetal alcohol syndrome (FAS) and fetal alcohol effects (FAE): Final report to the Centers for Disease control and Prevention. Seattle: University of Washington, Fetal Alcohol and Drug Unit.;1996.
23. Clarren S, Lutke J, Sherbuck M. The Canadian Guidelines and the interdisciplinary clinical capacity of Canada to diagnose fetal alcohol spectrum disorder. J Dev Behav Pediatr 2011;18:e494-e499.
24. Silver E, Smith W, Banks S. Constructing actuarial devices for predicting recidivism: A comparison of methods. Crim Justice and Behav 2000;27:733-764.
25. Moffit TE, Caspi A, Dickson N, Silva P, Stanton W. Childhood onset versus adolescent onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. Dev Psychopathology 1996;8:399-424.
26. Ruchkin V, Koposov R, Vermeiren R, Schwab-Stone M. Psychopathology and age at onset of conduct problems in juvenile delinquents. J Clin Psychiat 2003;64:913-920.
27. Castro Y, Carbonell JL, Anestis J. The influence of gender role on the prediction of antisocial bahaviour and somatization. International Journal of Social Psychiatry 2011;58**:**409-416.
28. Berkout OV, Young JN, Gross AM. Mean girls and bad boys: Recent research on gender differences in conduct disorder. Aggression and Violent Behavior 2011;16(6):503-511.
29. Maughan B, Rowe R, Messer J, Goodman R, Meltzer H. Conduct disorder and oppositional defiant disorder in a national sample: developmental epidemiology. Journal of Child Psychology and Psychiatry 2004;45(3): 609-621.