Abstract: International Neuropsychological Study of FASD

The long-term objective of this proposal is to determine if there is a unique neurobehavioral profile in children exposed to alcohol prenatally. Even though a large body literature exists on cognitive-emotional functioning in alcohol-exposed children, it is unknown if there is a signature neurobehavioral profile in children with fetal alcohol syndrome (FAS) or Fetal Alcohol Spectrum Disorder (FASD). Specific aims of the current investigation are: 1.to assess cognitive-emotional functioning in children diagnosed with FASD from a community in South Africa and a number of American Indian reservations in the Northern Plains States. A core test battery designed by the NIAAA-supported Consortium of International Collaborative Research on FASD will be utilized in this international study; 2. to test a specific statistical model of neurocognitive functioning (radex model) with the aim of further elucidating cognitive dysfunction in alcohol-affected children. The health relatedness of the project is that the identification of neurobehavioral profile in children with prenatal alcohol exposure will help clinicians diagnose children with FASD, specifically those alcohol-exposed children without evidence of dysmorphia. Identification of a profile of strengths and weaknesses of cognitive-emotional functioning is also essential for developing evidence-based interventions for children with FASD. Specifically, clinicians will be able to design programs that capitalize on strengths to address weaknesses in cognitive emotional functioning of these children. The research design of the proposed project involves a comparison of test performance of alcohol-affected children with that of typically developing children matched for age, sex, SES, and ethnicity. Potential participants in South Africa will be 150 children previously diagnosed as having FAS by our epidemiologic studies utilizing an international team of dysmorphologists. Case controls (n=150) will also be studied. The American Indian sample volunteers from a large group of children diagnosed with a FAS and other FASD (n=100) through the University of New Mexico FAS Epidemiology Research and Prevention Project. Case controls will be selected from the same communities. The test battery will be comprised of tests designed to measure general cognitive ability, attention, executive functioning, language, visual perception, memory, motor skills and emotional functioning. The data gathered through the core test battery will be combined with those collected by other participants in the international collaborative project (e.g. Russia, Finland, Italy, and other various sites in the US) to create a large data base. It is expected that this international collaborative project will provide a rare opportunity to researchers to answer many questions pertaining to cognitive-emotional functioning of alcohol-affected children to develop a behavioral phenotype and to design evidence-based interventions.

1. SPECIFIC AIMS:

Although there exists a large body of literature on the cognitive-emotional functioning in children with prenatal alcohol exposure, a neurobehavioral profile that is unique to alcohol-exposed children has not yet been delineated. Identification of a signature profile of a disorder is important both for improving diagnosis and developing evidence-based interventions.

The lack of progress in delineating a cognitive profile in children exposed to alcohol can be attributed to a number of factors. First, there has been difficulty in clearly defining the affected group due to inconsistent use of diagnostic criteria (Aase,1994; IOM,1996). Particularly, dysmorphologists disagree on diagnosing mildly affected children, creating a significant problem in subject selection for neurobehavioral studies (IOM,1996 ; Abel,1998). Second, researchers have used different instruments, sometimes of questionable validity and reliability, to assess a given construct (IOM,1996). Third, much of the published research is based on small samples, which limits the reliability of the results.

The documentation of high prevalence rates of Fetal Alcohol Spectrum Disorder (FASD) in some American Indian communities and in select non-western countries has dramatically changed the 'environment' for FASD research (May, et al.,1983; May,1991; May,McCloskey & Gossage,2002). Through international collaboration, researchers now have access to large samples of children with substantial prenatal alcohol exposure and subjects with consistently severe levels of FASD (Riley, et al.,2003). Issues related to reliability and validity of diagnosis can be resolved by having experienced dsymorphologists systematically evaluate affected children (May, et al.,2000). Furthermore, experienced researchers are now able to construct and validate test batteries that can be used cross-culturally (Adnams, et al.,2001), thus affording an opportunity to define a cognitive profile associated with prenatal alcohol exposure. Furthermore, researchers are able to capitalize on the known patterns of cognitive difficulties in alcohol-affected children to conduct hypothesis or theory-driven research.

The long-term goal of the current research is to delineate cognitive and emotional profiles of children exposed to substantial amounts of alcohol prenatally and with valid and substantial symptoms of FASD. Specific aims are:

1. to administer a neurobehavioral core test battery to children with confirmed prenatal alcohol exposure from a community in South Africa and on a number of American Indian reservations in the US. The data gathered through this test battery will eventually be combined with those collected at other international research sites

2. to test a specific statistical model of neurocognitive functioning (e.g. radex model) in children with prenatal alcohol exposure. Radex or hierarchical models of cognitive abilities posit that complex tasks that are at the top of the hierarchy load on what is known as general or 'g' factor. We hypothesize that those complex tests at the top of the hierarchy highly discriminate between children with substantial prenatal alcohol exposure and normal controls.

B. BACKGROUND AND SIGNIFICANCE:

B.1. Behavioral phenotypes and cognitive profiles

Nyhan (1972) introduced the term, behavioral phenotype, to denote a pattern of behaviors (e.g., aggressive self-mutilation) that characterizes the Lesch-Nyhan syndrome. In Nyhan's view, the organic genetic disorder gives rise to a characteristic pattern of behaviors through biochemical or neurophysiological or other biological mechanisms. Thus, the essence of this view is that the organic genetic disease determines its behavioral manifestations. In view of the fact that environmental and developmental factors also influence behaviors of a child with an organic genetic disorder, the Society for the Study of Behavioral Phenotypes adopted a broader definition of behavioral phenotype. According to this definition only some of the behaviors exhibited by children with biologically-based handicapping disorders are organically determined (SSBP, 1990). Furthermore, Plomin (1991) broadened the scope of the term behavior to include a range of observations including mental illness, learning skills, and intelligence. These strands of thought were later brought together in the following formal definition proposed by O'Brien and Yule (1995): "*The behavioral phenotype is a characteristic pattern of motor, cognitive, linguistic, and social abnormalities which is consistently associated with a biological disorder"* (pp. 2).

Abnormalities associated with a number of neurobehavioral disorders are often accompanied by dysmorphia and medical conditions. The behavioral phenotype of Lesch-Nyhan syndrome is often notable for compulsive self-injurious behaviors (e.g. finger and lip biting). Patients with Angelman syndrome tend to be very sociable and exhibit a distinctive pattern of inappropriate laughter (Smith et al. 1996). These patients show relatively better visual than verbal skills. A triad of impairments characterizes autism: qualitative impairments of social interaction, communication deficits and the presence of ritualistic/stereotyped behaviors and interests (Lord & Rutter, 1994). Children with autism perform remarkably well on selected visual spatial tasks such as creating designs with blocks and finding hidden figures (embedded figures), whereas they perform in the deficient range on tests of social-emotional reasoning (Shah & Frith, 1993). In contrast, children with Williams syndrome display an unusual command of language in the presence of impaired visual constructional skills (Bellugi, Wang, & Jernigan, 1994). In view of these findings, it is reasonable to ask if children with fetal alcohol syndrome (FAS) display a characteristic pattern of motor, cognitive, linguistic, and social abnormalities. Similarly, how do these characteristics translate to other levels of FASD?

B.2 Cognitive-social functioning in children with prenatal alcohol exposure

Intellectual functioning: Numerous researchers have consistently found intellectual deficits in children with fetal alcohol syndrome (FAS), with the average IQs of these children falling in the mild retardation to borderline range. Mattson and colleagues (1997) found that the mean Verbal and Performance IQ scores of children diagnosed with FAS were in the borderline range. A number of other researchers have reported that children with FAS, on average, performed in the mildly retarded range (Abel, 1990; Streissguth, Randels, &

Smith, 1991). Given that subtest scores provide useful information pertinent to defining behavioral phenotypes (e.g. splinter skills), researchers have conducted “profile analyses” of subtest scores from standardized test batteries (Adnams et al. 2001; Mattson et al. 1997). Results from these analyses show that children with FAS earn lower scores than case controls on all subtests. Having reviewed the literature on intellectual performance of alcohol-exposed children, Mattson and Riley (1998) also concluded that there was no significant discrepancy between scores measuring verbal and non-verbal (performance) abilities.

Attention and executive control functioning: Although there is a growing consensus among researchers that attention and executive control functioning are ‘core deficits’ in children with FAS spectrum disorder, there is less agreement among them as to what components of these skills are mostly deficient. Streissguth and colleagues (1998) found that adolescents with prenatal alcohol exposure were more impaired at tests assessing ‘focus’ and ‘sustain’ elements than at those assessing ‘encode’ and ‘shift’ elements of attention. In contrast, Coles and colleagues (1997) found that dysmorphic children with prenatal alcohol exposure were impaired only at shift and encode elements. In a subsequent study of sustained attention in alcohol-affected children, Coles et al. (2002) found, however, a significant modality x group interaction. That is, the alcohol-exposed group had more difficulty with sustaining attention when processing visual stimuli than when processing auditory stimuli.

Kodituwakku and colleagues (Kodituwakku et al., 1995; Kodituwakku et al. 2001) reported that alcohol-exposed children were markedly impaired with planning ability, as assessed by the Progressive Planning Test, a look-ahead puzzle. However, using the California Tower Test, Mattson el al. (1999) obtained less pronounced deficits in planning with alcohol-affected children. Most researchers have utilized the Wisconsin Card Sorting Test to assess the shifting aspect of attention. As noted above, researchers have found variable levels of impairments in the shift element of attention in alcohol-exposed children. Furthermore, results from parent or teacher-rated questionnaires often indicate that alcohol-exposed children are ‘impulsive’ and ‘stubborn’ (Streissguth et al. 1998). However, on most laboratory tasks that purport to measure response inhibition, children with prenatal alcohol exposure do not evidence impulsivity (Kodituwakku et al., 2001; Coles et al., 1997).

Language: Studies of language skills in children with prenatal alcohol exposure have also yielded inconsistent results. Numerous case studies report reduced language acquisition in children with FAS spectrum disorder (Abel, 1990). A number of group studies also have documented deficient language skills in alcohol-affected children. Becker, Warr-Leeper, & Leeper (1990) reported that children with FAS exhibited deficits in articulation abilities, comprehension, and the use of grammatical markers in repetition and spontaneous speech. Mattson and colleagues (1998) found that children with substantial prenatal alcohol exposure were impaired at tests of naming and comprehension. One of our studies conducted in South Africa revealed that the language subtest of the Griffiths Scales highly discriminated the FAS group from controls (May, et al.,2000;Adnams, et a., 2001). In some large prospective studies, researchers have failed to find a significant association between prenatal alcohol exposure and language development. For example, Greene et al. (1990)

did not find a significant relationship in the Cleveland cohort between prenatal alcohol exposure and measures of language development from the Sequenced Inventory of Communication Development. It should be noted, however, that this cohort included only one child with significant dysmorphia.

Learning and memory: There is consistent evidence that children with prenatal alcohol exposure perform less competently than controls on those learning and memory tasks that involve conscious effort for encoding and retrieval of information. Mattson and colleagues (Mattson et al., 1996; Mattson & Roebuck, 2002) have reported that children with heavy alcohol exposure obtained lower scores than controls on verbal and non-verbal list learning tasks such as the California Verbal Learning Test for Children. Verbal memory deficits in children with prenatal alcohol exposure have also been documented with other methods of testing such as story recall (Streissguth et al., 1989). There exist numerous reports in the literature on spatial learning, specifically on place learning in alcohol-exposed children. Using a spatial recall task, Uecker and Nadel (1996) found a general spatial memory deficit in alcohol-exposed children. Streissguth et al. (1994) have found that prenatal alcohol exposure was associated with deficient performance on a visually guided maze-learning test. Place learning difficulties in alcohol-exposed children have been recently demonstrated with a virtual Morris water task (Hamilton, Kodituwakku, Sutherland, & Savage, in press). The computerized test simulates the spacial memory dynamics of the Morris water maze test used for animals. There is also accumulating evidence that alcohol exposure does not produce impairments in recognition memory and procedural memory (Carmichael Olson et al., 1998; Mattson and Riley, 1999).

Motor: A few researchers failed to find an association between prenatal alcohol exposure and motor development (Chandler et al., 1996; Fried & Watkinson, 1990). However, there exist numerous neuropsychological and experimental studies of motor performance of alcohol-exposed children showing deficient manual dexterity and disturbance of balance (Roebuck et al., 1998; Mattson et al., 1998).

Social: Studies of social functioning of alcohol-exposed children have also yielded inconsistent results. Coles et al. (1991) evaluated children exposed to moderate amounts of alcohol during pregnancy (average age 6) and found no significant deficits in socialization skills as measured by standardized test of adaptive behavior. In sharp contrast, using the Vineland Adaptive Behavior Scale, Thomas et al. (1998) found that children with FAS were significantly more impaired in socialization skills than developmentally delayed children. In a recent study, Whaley, O'Connor, & Gunderson (2001) reported that in comparison to a clinical group with psychiatric diagnoses, those with prenatal alcohol exposure evidenced a rapid decline in socialization standard scores with age.

B.3. Sources of inconsistencies and potential solutions:

Thus, neurobehavioral research on FASD has produced somewhat inconsistent results, which can be attributed to some methodological issues. First, there is considerable variability in the populations included in the foregoing studies because of the lack of

uniformity in diagnosis and inclusion criteria (Riley, et al.,2003). While studies conducted in clinical or epidemiology research settings have included children identified primarily through dysmorphology evaluations, large-scale non-clinical studies have included children based on maternal reports of drinking during pregnancy (IOM,1996, Ch.5; Abel,1998). One cannot define a behavioral phenotype associated with a syndrome in the absence of a clear definition of the syndrome. Ideally, the syndrome should be defined, at least partially, independent of the behavioral phenotype. Accordingly, one can cogently argue that the neurobehavioral profile of those children who display dysmorphia (e.g. characteristics of the syndrome other than behavioral phenotype) should be established first. However, until recently, neurobehavioral researchers gained access only to a limited number of children with FAS or partial FAS. Second, researchers have used tests and procedures of variable sensitivity to assess cognitive-behavioral functioning of alcohol-affected children. For example, researchers have utilized continuous performance tests of variable complexity to measure sustained attention of alcohol-exposed children (e.g. NIMH battery, TOVA) and obtained inconsistent results. There is a growing consensus among researchers that narrow-band tests tapping specific constructs are more useful in defining behavioral phenotypes than standardized multi-factorial tests (Bellugi et al., 1994). Researchers have often used standardized tests such as IQ tests and achievement tests in the assessment of cognitive abilities in children with prenatal alcohol exposure. Intelligence tests have limited value in unravelling the puzzle of a FAS or other FASD phenotype.

As noted above, through studies of active case ascertainment of FASD in the United States, South Africa, and Russia, researchers now have access to large samples of children with very substantial levels of prenatal alcohol exposure and clear symptoms of FASD. Issues related to reliability and validity of diagnosis is addressed by having experienced dsymorphologists systematically evaluate affected children using the same diagnostic approach and criteria (May, et al.,2000). These large population-based studies also place the behavioral phenotype and cultural differences into a larger epidemiological perspective allowing more accurate interpretation of results of all domains (especially maternal factors of exposure and neuropsychology). Furthermore, a multi-site study will allow using a well-designed test battery on a very large sample of children with FASD drawn from different cultures. This will afford a rare opportunity to test hypotheses regarding cultural influences on test performance, which will eventually help define a behavioral phenotype directly associated with prenatal alcohol exposure. Furthermore, researchers are able to capitalize on the known patterns of cognitive difficulties in accurately and consistently diagnosed FASD children to conduct hypothesis-driven research.

B.4. A hierarchical model of ability organization and fetal alcohol syndrome

Many theorists have proposed a hierarchical model of ability organization (Cattell, 1971; Snow, 1978). At the top of the hierarchy is a general factor- or g factor- that accounts for performance on a wide range of tasks. The correlation between a test and g reflects the complexity of its required cognitive operations. Thus, more complex tests such as the Raven Matrices correlate highly with the g factor whereas relatively simple tests (visual memory) show low correlations with g (Marshalek, Loman, & Snow, 1983). Using a non-

metric multidimensional scaling method Guttman (1954) demonstrated that a test could be represented as a point in a two dimensional space defined by complexity and content (verbal, spatial, numerical). Thus, tests of varying complexity from different content areas can be represented in a circle called radex. Complex tests loading on g fall in the middle of the circle, while simple tests showing low correlations with g scale in the periphery. Marshalek et al. (1983) concluded that complex tasks might require more involvement of executive functioning.

Despite inconsistencies in reported research on FASD, there is now sufficient evidence that children with prenatal alcohol exposure are markedly impaired in relatively complex tasks, irrespective of the domain of functioning, whereas these children perform in the normal range on relatively simple tests. Tests of executive control functioning typify complex tasks that require considerable cognitive efforts. A number of researchers have reported deficient cognitive planning ability, as measured by look-ahead puzzles, in alcohol-affected children (Kodituwakku et al., 1995; Mattson et al. 1999). Alcohol-exposed children, however, are able to solve relatively simple look-ahead puzzles as competently as controls. There is consistent evidence that children with prenatal alcohol exposure are impaired at conceptual and affective set shifting (Coles et al. 1997; Kodituwakku et al. 2001). Researchers have obtained evidence that alcohol-affected children are deficient in letter fluency as indexed by difficulty in generating words beginning with given letters under specific constraints imposed by the examiner (Schonfeld et al. 2001). In contrast, performance of alcohol-affected children on tests of category fluency (generating exemplars of semantic categories such as animals, fruits and vegetables) is unimpaired (Kodituwakku et al. 1995). The letter fluency task is more demanding than the category fluency task in that the former requires more executive control than the latter. Children with prenatal alcohol exposure demonstrate deficient free recall of information although their recognition memory is relatively intact. Again, free recall of information is more demanding than recognition of information. Furthermore, children with prenatal alcohol exposure are markedly impaired in tests of fluid intelligence such as the Raven Matrices (Kodituwakku et al. 1995). Thus, it is reasonable to hypothesize that complex tests that load on g (tests that scale in the middle of radex) are more sensitive to prenatal alcohol exposure than relatively simple tests that scale in the periphery of radex.

Functional neuroimaging studies have revealed that complex tasks that load on g, such as the Raven Matrices or the Tower of London task, activate multiple brain regions in normal adults including the bilateral prefrontal cortex and left parietal lobe. Therefore, one plausible hypothesis is that alcohol-exposed children have a deficit in the ability to recruit different regions of the brain necessary for the execution of complex tasks. This hypothesis is consistent with the deficient neural integration hypothesis proposed for explaining some neurodevelopmental disorders (e.g. schizophrenia).

B.5. Significance:

Delineation of the behavioral phenotype in children with substantial prenatal alcohol exposure will allow clinicians to accurately diagnose children with the full range of FASD, specifically those at the mild end of the spectrum. A key recommendation of the Institute of

Medicine committee for improving diagnosis was, “investigation of the differences in expression and specificity of behavioral deficits in FAS and ARND” (IOm,1996,p. 6). It is now documented that those children who do not receive services because of a lack of diagnosis are more likely to develop secondary disabilities during their adolescence and adulthood (Striessguth et al., 1996). Furthermore, the identification of a profile of strengths and weaknesses in alcohol-exposed children is a critical step toward developing evidence-based interventions for them.

C. PRELIMINARY STUDIES

C.1. General cognitive ability:

Our research team has conducted two studies of patterns of cognitive-motor development in children with FAS from a community in the Western Cape Province of South Africa (Adnams et al., 2001) and a third one is currently underway. In the first study, 34 children with FAS and 34 age (FAS mean age=6.99 years; control mean age= 7.12 years), gender, and ethnicity matched controls participated. The Griffiths Mental Developmental Scales, normed for the South African population, were utilized to assess cognitive-motor development. As Figure 1 shows, four Griffiths scales discriminated highly the FAS group from controls: hearing and speech (language), eye-hand coordination (fine-motor), performance (pattern construction), and practical reasoning. In sharp contrast, no significant group differences found on the gross motor and personal-social (simple adaptive skills) scales. It should be noted that the scales which were most highly discriminating between the two groups involve higher-levels of intellectual skills.

Figure 1: Griffiths mean scores for FAS and control children

Source: Adnams, et a., 2001.

There is converging evidence from several lines of research that support a distinction between fluid and crystallized intellectual abilities. Fluid intelligence refers to the ability to solve novel problems without relying on an established knowledge base acquired through experience and schooling. In contrast, crystallized intelligence relies on a previously acquired knowledge base. We tested the hypothesis that tests assessing fluid intelligence

are more sensitive to the effects of prenatal alcohol exposure than those measuring crystallized intelligence (Riley et al., in press). As part of a neuropsychological test battery, the Raven Progressive Matrices were utilized to assess fluid intelligence, and the Peabody Picture Vocabulary Test to assess crystallized intelligence. Twenty alcohol-exposed children (mean age= 10.80 years) and 20 controls (mean age = 10.65) participated. The two groups were also matched for gender and ethnicity. The results showed that the alcohol-exposed group performed significantly worse than controls on the Raven Matrices [t (38)= 3.63, p < .001], but that the two groups did not differ on the PPVT [t (38)=1.93, P = .06].

C.2. Executive control functioning:

Subsumed under the umbrella term, executive functioning (EF), are a number of abilities involved in achieving a goal in an efficient manner. These include planning, set shifting, and response inhibition. Evidence from both animal and human research points to two domains of EF: cognition-based and emotion-related (Dias et al.1996). Commonly used tests of EF, such as cognitive planning, conceptual set shifting, and rapid generation of verbal or non-verbal responses, assess cognition-based EF. Visual discrimination reversal and extinction tests are considered classic tests of emotion-related EF. We have conducted research aimed at assessing both domains in individuals exposed to alcohol prenatally.

Cognitive planning: Look-ahead puzzles are often used to measure cognitive planning abilities (Owens, 1996). We utilized a look-ahead puzzle called the Progressive Planning Test (PPT) to assess cognitive planning skills in children with substantial prenatal alcohol exposure (Kodituwakku et al., 1995). In this test, the participants are required to move 1 to 5 beads from an initial position to a goal position while following specific rules (See section D.3 for details). The test consists of 12 problems of increasing difficulty. The number of mental manipulations (mental reversals) one has to carry out to solve a problem determines its complexity. Thus, the problems on the test can be categorized into 3 sets, each containing 4 problems of comparable complexity. The problems in sets 1, 2 and 3 involve 0, 1, and 2 mental reversals (mental manipulations) respectively. We found that children with prenatal alcohol exposure solved fewer problems on this test than controls [t (18)=5.04, p< .001]. In a subsequent study, we obtained evidence that the group effect on planning remained highly significant after controlling for the group differences in intellectual ability [F(37, 1) = 11.902, p<.001]. In follow-up analyses, we found that only complex planning problems that involve mental manipulation discriminated the two groups. As shown in Figure 2, the alcohol-exposed group performed as competently as the control group on set 1, but the former performed markedly worse than the latter on set 2.

Figure 2: Differential proportions of successful performance of the alcohol-exposed children on level 1('easy') and level 2 ('complex') planning problems

Conceptual set shifting: Conceptual set shifting refers to the ability to shift attention from one dimension of a stimulus to another based on feedback. In a test most commonly used to assess this ability [Wisconsin Card Sorting Test (WCST)], the subject is asked to sort cards by a given dimension (e.g. color) and then to shift attention to sorting them by a different dimension (e.g. shape) according to the examiner's feedback. Thus, the WCST measures the subject's ability to shift attention across concepts or dimensions, hence called extra-dimensional set shifting. In a number of research projects, we have demonstrated that children with prenatal alcohol-exposure are deficient in conceptual set shifting, as measured by the WCST (Kodituwakku, et al., 1995; Kodituwakku et al. 2001). Children with prenatal alcohol exposure make significantly more perseverative errors on this test than controls [$t(38) = 2.89$, $p < .01$].

Letter and category fluency: Rapid generation of words beginning with a specific letter (e.g. F,A,S) under some constraints (e.g. excluding names of people or places), or letter fluency, involves EF. The letter fluency is a complex task in that it requires invoking a number of operations simultaneously: rapidly generating words by phonemic similarity, checking responses to ensure that they meet the test constraints, and keeping a record of responses already produced in working memory (to avoid repetition). The letter fluency can be contrasted with category fluency, in which subjects are required to rapidly generate exemplars from a semantic category (e.g. animals, fruit and vegetables). We reported that alcohol-affected children, compared to controls, are impaired in letter fluency, but unimpaired in category fluency (Kodituwakku, et al. 1995). In a recent study in South Africa we replicated this finding (manuscript in preparation).

Emotion-related executive control functioning: Emotion-related action selection relies on rewards and punishments. The ability to alter behavior in response to the changes in reinforcement conditions is considered to be a marker of emotion-related EF. We assessed emotion-related EF in alcohol-exposed children using a visual discrimination reversal test (Kodituwakku et al., 2001). In this test, the subject was required to respond to one of two patterns presented on the computer screen. One of the images was arbitrarily chosen as rewarding and the other non-rewarding. The subject was required to respond to the rewarding stimulus and withhold responding to the non-rewarding one. When the subject reached a learning criterion (9 correct in 10 consecutive responses), the reinforcement contingencies reversed without warning. If the learning criterion is reached again, further reversals occurred up to a maximum of 3 reversals. In the second part of the test, extinction

trials were administered. That is, after reaching the initial learning criterion, the subject was able to earn points only by withholding responses to both images. The results showed that the alcohol exposed children made fewer reversals than controls on the reversal learning condition [t (38) = 3.947, p < .0001].

Facial and vocal affect identification: The ability to identify facial and vocal emotional expressions (emotional prosody) is an important building block of pro-social behavior. We have conducted two studies of identification of facial and vocal emotional expressions in children with prenatal alcohol exposure respectively. In the first study on facial expressions, 40 children (20 alcohol-exposed and 20 controls), ages 7 through 18, participated. Selected subtests from the Florida Affect Battery were utilized to assess facial affect identification. The results showed that children with prenatal alcohol exposure were markedly deficient in facial affect identification [F (1,38)= 14.31, p< .001], especially on relatively complex subtests such as facial affect naming. Participants in the second study were 41 children (25 alcohol-exposed; 16 controls), who ranged in age from 7 to 15. As part of a neuropsychological test battery, 4 subtests from the Florida Affect Battery assessing emotional and non-emotional prosody were administered to the two groups. Data analyses revealed a significant group x subtest interaction [Wilks' Λ =.85, p < .05]. As Figure 3 shows, the alcohol-exposed group performed markedly worse than control on the conflicting prosody subtest, in which the participants were asked to name of the emotional tone ignoring the semantic content (e.g., The statement, 'all the puppies are dead' said in a happy voice). This is a demanding test in that the participants are required to inhibit their prepotent tendency for naming emotions based on the semantic content.

C.3 Memory:

On tests of memory and learning children with prenatal alcohol exposure demonstrate normal rates of acquisition and retention of relatively simple information, but they evidence difficulty in learning relatively complex information. In South Africa, our research team contrasted the performance of children with FAS and controls on the Lhermitte-Signoret

memory tests, which comprise two parts: spatial and logical (See Section D4, p.20 for details). In the spatial condition, the participants were required to learn the locations of 9 line-drawn objects. The logical memory condition involved learning a logical pattern of colored geometric shapes. The participants were 64 children diagnosed with FAS and 63 age, gender, and ethnicity matched controls. The results showed that the FAS group performed as competently as the control group in the spatial memory condition [$F(1,122) = 3.11$, $p = .08$], but that the former performed markedly worse than the latter in the logical memory condition [$F(1,122) = 7.32$, $p < .01$]. Similarly, on a test of recognition memory span, the FAS group displayed relatively intact recognition memory, but exhibited impaired free recall of information (data not yet published).

Thus, despite inconsistencies in the findings on neurocognitive functioning of children with FASD, there is an emerging pattern of results. That is, children with FASD tend to perform worse than controls on relatively complex tasks, irrespective of the domain of functioning. This pattern of results conforms to the radex or hierarchical model of cognitive abilities. Delineation of the topography of cognitive abilities in alcohol-exposed children is an essential step toward finding the neurobiological basis of cognitive dysfunction of these children. Functional neuroimaging studies of brain activity generated by complex tasks in normal subjects have revealed that those tasks recruit multiple regions in the brain simultaneously. For example, complex problems on the Raven Matrices activate bilateral frontal regions and the parietal regions, suggesting that successful performance on these problems probably requires a higher level of neural integration. Therefore, it is reasonable to hypothesize that children with prenatal alcohol exposure may have a deficit in neural integration, which depends on the integrity of both gray and white matter.

4. RESEARCH DESIGN AND METHODS:

D.1 Overview of research design:

The specific aims of the current study are to administer a core test battery to two large groups of children with varying levels of severe FASD as part of a multi-site international study and to test a specific statistical model of the topography of cognitive abilities in alcohol-exposed children. We propose to attain these specific research goals using a case-control research design, in which the performance of alcohol-exposed children will be compared with that of controls on a carefully selected test battery. The ability to draw valid conclusions within this research design hinges on the success of accounting for a multitude of variables that directly and interactively influence social-cognitive development in children with prenatal alcohol exposure. It is now established that children of alcoholic parents display temperamental and physiological differences when compared to those of non-alcoholic parents suggesting the influence of genetic factors (Begleiter & Kissin, 1995; Tarter, Moss, & Vanyukov, 1995). Furthermore, alcohol-exposed children often experience adverse conditions in their postnatal environments, including poverty, exposure to violence, and disruptions in attachments. Therefore, in the current research project we plan to compare alcohol-exposed children to a group of non-exposed children matched for a range of factors including socio-economic status, ethnicity, sex, age, and maternal drinking level and pattern (e.g., quantity, frequency and timing, binge vs. chronic, and peak BAC estimates). Information on family history of alcoholism, especially on paternal drinking

history, will also be ascertained to determine the comparability of the groups with respect to genetic influences.

D.2. Participants:

Potential participants in the current research will be from an existing pool of children with FAS from a community in the Western Cape Province of South Africa (N≈150), and an existing and growing pool of American Indian children from the Northern Plains States diagnosed as having FAS or partial FAS (N≈ 50) or alcohol related effects (ARBD and ARND; N≈ 50). Case controls for these participants will be selected from the same communities (N = 250) and matched as described above.

South African participants:

Wave I: Our team of US-South African researchers (Philip May, P.I.) has now completed three waves of NIAAA funded epidemiologic research involving screening for FAS in first grade classes in a community from the wine producing region of the Western Cape Province. The methodology utilized in these screening evaluations is described in detail by May and colleagues (May et al., 2000). In the wave I of screening conducted in 1997, a census of 406 children in first grade classes in six of the twelve local schools (urban and rural) were screened fully by dysmorphologists to establish local growth parameters and to diagnose possible FAS children. In the remaining six schools, only those first graders who fell below the 10^{th} percentile in head circumference or height and weight (N=220) underwent full morphology evaluations. To determine the reliability of diagnoses, two teams of dysmorphologists, who were blinded to the child's alcohol exposure history and cognitive functioning, evaluated each child. Inter-rater reliability of particular diagnostic features were found to exceed .90 (r) in most all physical measurements, .80 in a few others, but less than .80 in one, palpebral fissure length (.64). Therefore, reliability of the diagnosis was established for single examiners and ensured on all measurements by the use of multiple examiners. In Wave I, 53 children received a preliminary diagnosis of FAS or Deferred based on morphology examinations conducted in the above two phases. Matched case controls (on age, sex, and residence) were selected from among those who did not meet the above screening criteria. Neurobehavioral tests were administered to the two groups by a test administrator who was blinded to the morphology and alcohol-exposure data. Interviewers, who were also blinded to the morphology and neurobehavioral data, conducted interviews with mothers to ascertain information on prenatal alcohol exposure of these children (Viljoen, et al.,2002). Forty-six of the preliminarily deferred children received a final diagnosis of FAS, based on growth and morphology data, exposure information, and neurobehavioral test results. The rate of FAS in this study was 46 per 1,000, the highest ever reported in a functioning community anywhere in the world.

Wave II: In the second and third waves of screening, which were conducted in 1999 and 2002 respectively, the same two-tiered approach to diagnosis was utilized. In 1999, a total of 863 first grade students had tier 1 screening (height, weight, head circumference) and 399 of these children underwent tier 2 examination (full morphology) along with 120 randomly selected controls for the epidemiology study. Again, two teams of physicians

examined each child. Based on morphology data, 92 children were assigned a preliminary diagnosis of FAS or deferred (possible FAS). A South African psychologist (a native Afrikaans speaker) administered a battery of neurobehavioral tests to these 92 children and controls. Maternal interviewers (also South African) conducted interviews in Afrikaans with mothers to ascertain information on drinking variables (quantity, frequency, and timing of drinking). Final diagnoses were determined in a case conference after weighing results from all three domains: morphology and growth data, maternal interviews, and neurobehavioral test results. Sixty-four of 92 children were assigned a final diagnosis of FAS. This produced the highest FAS rate in the world to date: 75 per 1,000 children (results yet published, but to be submitted summer, 2003)

Wave III: In 2002, a total of 869 first grade children were screened and 447 of them were referred for a full morphology evaluation (including 134 randomly selected controls for the epidemiology control groups). Based on morphology data, a preliminary diagnosis of FAS or possible FAS was assigned to 108 children. Brief neuropsychological examinations and maternal interviews are currently in progress. A case conference to determine final diagnoses has been scheduled for September of 2003. We expect that 70 to 75 of the 108 children may receive a diagnosis of FAS following the case conference. We may also begin to utilize the other IOM categories of FASD in this wave, particularly the category of Partial FAS.

Thus, a minimum of 150 children with FAS will have been diagnosed in the above NIAAA-funded three waves of screening and most will be available for the current research project. Demographic information such as age and sex of these potential participants is presented in Table 1. Parents of these children are predominantly farm workers who are of mixed ethnic heritage known as 'Coloured'. The control group will be 150 children who do not evidence any dysmorphia consistent with FAS or other alcohol related birth defects and whose mothers do not have a history of moderate or heavy levels of prenatal alcohol exposure. Furthermore, the neuropsychological control group will be comparable to the FAS group with respect to SES, ethnicity, age, and sex.

American Indian participants:

As part of an ongoing NIAAA-funded prevention trial of fetal alcohol syndrome in the Northern Plains states (UO1 AA11685 - Philip May, P.I.), our research team has thus far screened 498 American Indian children with suspected histories of prenatal alcohol exposure. Using active case ascertainment methods, the research team continues to hold clinics on six reservations: Turtle Mountain Chippewa (North Dakota), Cheyenne River Sioux (South Dakota), Sisseton-Wahpeton Sioux (South Dakota), Flathead Confederated Salish and Kootenai (Montana), Ft. Belknap- Assiniboine and Gros Ventre (Montana), and The Blackfeet Nation(Montana). Also, one urban site has been added as a source of referrals and cases, Great Falls, Montana. We have proposed in our competitive renewal of UO1 AA11685, that all six of these reservations and Great Falls continue as sites for FASD research for five more years.

Table 1: Children Diagnosed with FAS and FASD in the Plains U.S.A.and Western Cape, South Africa

Plains Children

Age	Children w/FASD*	Male	Female
Birth - 3 yr.	**8**	**6**	**2**
3 yr. – 6 yr.	**16**	**8**	**8**
7 yr. – 14 yr.	**37**	**23**	**14**
TOTALS	***61***	***37***	***24***

*** The 61 children listed as subjects are drawn from a total sample of 498 Plains Indian children seen in the project through March, 2003. These children were diagnosed from the 85 most severe cases seen to date and are the only ones formally case-conferenced thus far. Therefore, a first diagnosis of some level of FASD for the remaining 413 children awaits a final case conference. We can conservatively estimate that at least 100 additional children in this existing sample will meet IOM diagnostic criteria for ARND, ARBD or both ARND/ARBD. Each of these estimated 100 children will provide possible subjects for this study. Additionally, we are scheduled for eight additional diagnostic, referral-based clinics in the remainder of 2003, and even more for 2004. On average, we examine and assess 20 to 25 children per one week clinic.**

South African Children

Wave	Age*	Children w/FAS	Male	Female
SA I (1997)	**6 yr. – 9 yr.**	**46**	**50**	**45**
SA II (1999)	**6 yr. – 8 yr.**	**64**	**32**	**37**
SA III (2002)	**6 yr. – 8 yr.**	**75****	**38*****	**37**
	TOTALS	***185***	***93***	***92***

*** The children are listed by age at time of initial diagnosis. Since the children in Wave I diagnosed with FAS were between 6-9 years in 1997 ($\bar{x}$ = 7.0 years) and in Wave II in 1999 ($\bar{x}$ = 6.5 years), these children are now between 12 to 15 years ($\bar{x}$ = 13) and 10-12 years ($\bar{x}$ = 10.5)**
**** 107 children in Wave III have a preliminary diagnosis of FAS or Deferred. In Wave II, 69.9% of the children with these preliminary diagnoses received a final diagnosis of FAS. Therefore, weconservatively estimate that 75 of the 107 subjects from Wave III are likely to have FAS.**
***** Sex for Wave III estimated from combined experience in Waves I and II.**

In the Plains, as in other active case ascertainment studies in the United States, clinical evaluations of the children are completed in a two-step process (May, et a., 1983; May and Hymbaugh, 1982; May and Gossage, 2002). First, children with a suspected history of prenatal alcohol exposure are referred by full time field staff at each site and examined by two experienced clinical geneticists/dysmorphologists (Eugene Hoyme, M.D. and Luther Robinson, M.D.). An experienced educational diagnostician (Wendy Kalberg, M.A., C.E.D.) administers a brief test battery aimed at assessing IQs and adaptive skills of these children. An experienced maternal interviewer (Phyliis Trujillo) conducts interviews with available mothers (N=101) to ascertain extensive information on maternal risk factors and quantity,frequency, and timing of drinking, nutrition, SES, gravidity, parity and peak BAC. Second, those children with dysmorphia and/or a history of substantial alcohol exposure are further evaluated with an extensive neuropsychological test battery. Based on morphology, maternal risk factors, and neurobehavioral data, final diagnoses of the children are determined in a case conference.

Of the children who have gone through the above process, 48 have received a diagnosis of FAS or partial FAS. Five children have been diagnosed as having alcohol-related birth defects (ARBD), and 8 as having alcohol-related neurodevelopmental disorder (ARND); Diagnosis of 8 other children has

been deferred on the grounds of insufficient data at the time of case conference (most were too young [<2yrs] at the time of exam). A cross-tabulation of these children by sex and age is presented in Table 2 (page 25). The number of potential participants in the current research over the next 5 years can be conservatively estimated at 200 (100 FAS spectrum disorder; 100 controls).

Recruitment:

In the South African study site, a social worker in the community or the child's classroom teacher will contact the parents of children with FAS and controls and explain the study. A significant number of parents in the South African community are marginally literate or illiterate. In the Native American communities, research site coordinators currently employed by UO1 AA 11685 (FAS Epidemiology Research and Prevention) will contact parents/ children through a letter that explains the study. The control group will be recruited through advertisements and announcements in tribal centers and community gatherings. The following exclusionary criteria will be utilized in the recruitment of the potential participants: moderate or severe mental retardation, head trauma, co-morbid genetic syndromes, severe psychiatric illness (psychosis), substance abuse, and neurological problems such as seizure disorder.

In both sites, our colleagues have extensive experience in slecting control groups. Control groups are matched with the FAS group with respect to age, sex, ethnicity, and SESin both the American Indian communities and in South Africa.

D.3. Specific Aim 1:

The first specific aim of the current project is to administer a core neurobehavioral test battery to children with confirmed prenatal alcohol exposure from a community in South Africa and on a number of American Indian reservations in the US. The data gathered through this test battery will eventually be combined with those collected at other international research sites, including Russia, Finland, Italy and the US. Thus, the current research helps create a large data set and affords researchers a rare opportunity to test the effects of prenatal alcohol exposure and environmental variables on cognitive dysfunction in the affected children. If alcohol-exposed children evidence a comparable profile of test results despite cultural variability, one can strongly argue for a behavioral phenotype associated with prenatal alcohol exposure.

Testing Procedures:
The following core test battery recommended by a group of experienced researchers for use in international FAS research will be utilized to attain specific aim 1.

General cognitive ability:

Leiter International Performance Scale- Revised (Stoelting, 2001)

This measure provides an assessment of global ability through its Visualization and Reasoning (VR) Domain battery. There is a screening IQ assessment that can be administered in 25 minutes. The Leiter-R is correlated with other frequently used measures of ability, including the WISC-III (.85). It also includes "growth" scores that can be used to assess progress in children with developmental disability and to chart relatively small difference in cognitive performance over time or between groups.

The Leiter-R has several characteristics that make it appropriate for the collaborative study. It is completely nonverbal and does not require a spoken word from either examiner or child. The child

is not required to read or write any material. For this reason, the same form can be used in different countries and in different cultural groups. In fact, the Leiter-R is widely used in Europe. Apparently, the manual has been translated into Italian and German. (For this study, translation in to the remaining non-English languages used in the study (e.g. Afrikaans) would be required.) The norms are based on a United States population but included Native American, Hispanic, Asian-American and Hispanic groups as well as Caucasian. The test manual notes that Leiter-R standard scores in these groups were less disparate from those of the majority culture than were scores on the WISC-III.

Attention and executive functioning:

Test of Variables of Attention (TOVA; .Greenberg, 1993)
The TOVA is a computerized continuous performance test that assesses processing abilities in both the visual and auditory modalities. Attention and impulse control are measured in the following three areas: inattention scored by omissions, impulse control scored by commissions, and response time variability.

The visual TOVA provides age norms by year based on a sample of 1,340 males and females ages 4 through 19 years. The task presents the subject with either a target or non-target geometric picture centered on the computer screen for 100 msec every 2 s. The stimuli are colored squares with smaller black embedded squares. For the target stimulus, the embedded black square is at the top of the colored square. For distracter stimuli, the embedded black square is at the bottom of the colored square. The auditory TOVA task presents the subject with either a target tone ("Middle G") or a non-target tone ("Middle C") for 100 msec every 2 s. The auditory output is presented through external computer speakers. The goal of each task is to correctly respond to each target by pressing a button as quickly as possible. During the first half of the test (stimulus infrequent condition), the designated target is presented 22.5% of the time. During the second half of the test (stimulus frequent condition), the target stimulus is presented 77.5% of the time. The function of the different conditions serves to examine the effects of differing demands on attention and inhibition. The presentation of test stimuli is broken down into four quarters, two quarters per half. Stimuli are presented at a fixed random frequency per quarter. Performance scores are determined per quarter, per half, and in total. Each quarter lasts a period of 5.4 minutes, each half lasts 10.8 minutes, and the entire test lasts 21.6 minutes. In addition, a 2.25 minute practice test is administered at the beginning of each test session.

Different variables are measured throughout the test and the four primary variables will be discussed here. Errors of omission are interpreted as a measure of inattention and are marked by a failure to respond to the designated target. Errors of commission are interpreted as a measure of impulsivity or disinhibition and are marked by an incorrect response to a nontarget. Response time is a measure of processing time for a correct response to the target. Variability measures the subject's inconsistency in response times and can be viewed as a variance score, reported in terms of standard deviations of the mean correct response times. The software automatically records the subject's responses and calculates standard scores and *z* scores for each variable.

Cambridge Neuropsychological Test Automated Battery (Morris et al., 1987).

The Cambridge Neuropsychological Test Automated Battery (CANTAB) was selected as an instrument to assess visual attention due to its ability to reliably assess cognitive functioning and underlying neurological mechanisms. The subtests of the CANTAB were designed to assess a range of abilities subserved by the frontostriatal circuitry, including planning, set-shifting and spatial memory. Originally developed to assess cognitive function in elderly and demented patients, the

CANTAB test battery has subsequently been used with other clinical populations and children.

The CANTAB utilizes touch screen technology to administer the items and record responses in a number of its tests. This allows testing with limited verbal requirements and provides accurate measurement of reaction time (RT). Testing will be conducted on an Advantech PPC 120 computer with a 12-inch touch-sensitive screen. As suggested in the CANTAB administration manual, all participants will be seated 60 cm away from the screen to discourage resting a hand on the computer and will be administered the same subtests in the same order. The following subtests from the CANTAB test battery will be utilized to determine the presence of requisite basic skills to perform the tests and to assess skills in attention and executive functioning.

Motor Screening (MOT). This test is primarily used as a screening task and is administered at the beginning of all CANTAB batteries to ensure that the subject can identify and touch the targets presented on the screen. In this task, crosses are displayed on the screen and the subject is directed to touch the cross with the forefinger of the dominant hand. This test provides a baseline measure of response latency (msec) that can be used to compare with same-subject performance on subsequent CANTAB tests.

Big-Little Circle (BLC). This test is designed to assess the subject's ability to follow an explicit rule, and then reverse the rule. In this task, the subject is presented with two boxes on the screen: one containing a big circle and the other containing a little circle. The subject is instructed to touch the little circle. After 20 pairs, the rule is changed, and the subject is instructed to touch the big circle. Dependent variables are the percent correct and mean latency (msec) for correct and incorrect responses.

Delayed Matching to Sample (DMS). This test assesses the ability to remember a stimulus and then match it to the same stimulus among distracters. Thus it involves a type of *what* visuospatial processing and taps temporal function. The target is a complex visual stimulus that varies in color and shape. The target stimulus is presented briefly, followed by four choices, one of which is the correct response. The three distracters vary along the dimensions of color and shape. There is a variable delay between target and response choices of either 0, 4, or 12 seconds. In addition, there are "simultaneous" trials in which the target and the response choices appear concurrently. There are three practice trials followed by 40 test trials (10 each of the simultaneous, 0, 4, and 12 second conditions). The dependent measure is the percent correct for each condition.

Spatial Span (SSP). This task is a measure of visual memory, similar to the Corsi block test, requiring the subject to remember the sequence in which squares on the screen change color. Thus it represents a measure of visual *where* memory, and taps parietal function. Trials begin with two squares changing color, one at a time and the participant is instructed to touch the same squares in the same order. There are three trials at each span length and span lengths increase by one until all three trials at a given length are failed or until the maximum span of nine is reached. There are two dependent measures: spatial span (maximum sequence correctly remembered), and an error score indicating the mean number of incorrect boxes selected per trial.

Intra-Extra Dimensional Shift (IED). This subtest is a measure of the subject's ability to attend to a specific perceptual feature of a stimulus, and then shift attention to the previously irrelevant dimension. Utilizing feedback from the computer, subjects are required to determine which of two dimensions (e.g., shape or a line) is relevant, and maintain that focus with different examples (intra-dimensional shift). After six successive correct responses are made, the relevant stimulus changes, requiring the subject to shift attention to the previously irrelevant dimension (extra-dimensional shift). Dependent variables are stage achieved (out of 9 possible), trials to criterion, and mean latency (msec) for each stage.

Stockings of Cambridge (SOC). This is a test of spatial planning, similar to the Tower of London test. In this task the subject sees two displays of three stacked colored balls. The object is to make the balls in the lower display match those of the upper display. Balls are moved one at a time by touching and dragging to a new location. This task begins simply (1 move) and increases in complexity up to 5 moves. Time to completion and number of moves required are recorded. This task also includes a control for motor performance. The upper display moves balls one at a time and the subject follows the pattern on the lower display. This allows for a comparison between the amount of time expended on motor activation (motor control task) and that required for planning (test task).

We also propose to use the following subtests from the Leiter –R Attention and Memory battery to supplement the above measures of attention and executive functioning: Attention Sustained, Attention Divided, Forward Memory Span, Reverse Memory Span, Spatial Memory Span, and Visual Coding.

Attention Sustained: Analogous to commonly used letter cancellation tasks, this test requires the child to rapidly search and cross out a shape that is randomly distributed among distractors on a page. The target shape is printed at the top of each page, thus reducing the memory demand of the test. The number of target shapes correctly identified within a given time limit is used as a measure of sustained attention. In the Mirsky's model of attention (Mirsky & Duncan, 2001), however, visual search also involves 'focus attention', which is considered to be mediated through the parietal lobes.

Attention Divided: In this test, the child is asked to do two things alternatively: sorting a number of cards and pointing to specific pictures. The number of cards sorted correctly and number of pictures pointed correctly are taken as an index of divided attention.

Memory Span: The following four subtests from the Leiter will supplement the CANTAB Spatial Span Test in the assessment of memory span. In the Mirsky's model of attention, span tests tap what is termed 'encode' element of attention.

Forward and Reverse Memory Spans: The Forward and Reverse Memory Span Tests are analogous to the Corsi Blocks Tapping Test. In the Forward Memory Span test, the examiner points to picture sequences of increasing length and the child is asked to point to them in the same order. In the Reverse Memory Span test, the child is asked to point to the sequences in the reverse order.

Spatial Memory Span: On a given trial of this test, the child is shown a grid with pictures printed on specific locations for 10 seconds. Then the picture grid is removed and the child is presented with a blank grid and the pictures that were displayed. The child's task is to place the pictures on the locations where they were. Increasing numbers of pictures are shown across trials to determine the spatial memory span.

Visual Coding: As in the Wechsler Digit-Symbol Test, the child is presented with a "key" at the top of each page showing a row of shapes paired with a row shapes or numbers. Shown below the key is the top row from key in which the shapes are rearranged. The child is asked to find the shape or the symbol from the bottom row that goes with each shape. Complex items on the test involve combining shapes to come up with a new shape, a task called code learning.

Learning and Memory: The following subtests from the CANTAB battery will be used to assess learning and Memory.

Pattern Recognition Memory (PRM) This is a test of visual, or *what* memory and involves the ability to discriminate between a previously displayed pattern and one that has not been displayed. In this task 12 different patterns appear in a box on the screen, followed by two boxes, one that displays a pattern from the earlier phase while the other box displays a novel pattern. The participant is instructed to touch the box containing the pattern that was previously seen. There are two sets of 12 trials (24 total trials). Dependent measures are accuracy (%correct patterns selected) and mean RT for correct and incorrect responses.

Spatial Recognition Memory (SRM) This is a test of spatial, or *where* memory and taps parietal function. In the learning stage of this task, five squares (targets) are presented one at a time in different locations on the computer screen. During the choice stage, one target is presented along with a distracter in a novel location. The participant is asked to indicate the target location by touching that square. This repeats for the remaining four boxes in the stimulus set. There are four trials of 5 boxes. Dependent measures are accuracy (%correct locations selected) and mean RT for correct and incorrect responses.

Paired Associates Learning (PAL) This test comprises eight trials in which the subject is required to remember both *what* the target pattern was and *where* it was presented in a series of squares on the screen. First, the patterns are revealed one at a time to the participant during the learning phase; during the test phase the participant is shown the same patterns and asked to touch the appropriate squares. If the participant responds correctly, a new learning phase begins. If the response is incorrect, then the previous learning phase repeats until the correct response is made. There are two trials of one, two, and three patterns, then one trial each of six and eight patterns. If a participant does not meet criteria for moving to the next trial after ten presentations, the test is automatically terminated. Two dependent measures will be used from this task: the number of trials completed and the number of errors.

Two subtests from the Leiter will be used to assess memory and learning: *Associated Pairs and Delayed Pairs.* In the Associated Pairs, a test of paired associate learning, the child is shown pairs of line drawn objects, the number of pairs on display varying from 2 to 8. During the test phase, the child is shown one drawing from a pair asking to find the other one.

In the Delayed Associate Pairs, memory of associate pairs is tested following a filled delay.

Visual-spatial and visual-constructional abilities:

Developmental Test of Visual Motor Integration (Beery, 1997)

The Beery Developmental Test of Visual Motor Integration (VMI) assesses visual-motor integration by requiring the child to copy 24 geometric figures [9]. This test is constructed to measure the level of visual-motor development from the age of three. Each design is harder than the last and testing is continued until three consecutive figures are failed. There are two supplemental tests: the VMI Visual Perception and the VMI Motor Coordination tests. These tests use the same stimulus forms and thus are comparable to the main test. Normative data are available for all ages and administration time is approximately eight minutes.

Road-Map Test of Direction Sense (Money, 1976)

In this test, the examiner traces a dotted pathway with 32 turns and the subject is required to tell the direction (left or right) taken at each turn. Some of the turns on the test are un-rotated (in the

direction the subject is facing) and the other turns are rotated. Developmental norms are available for ages 5 through 18. (Money, 1976)

Finger Localization Task (Benton et al, 1994)

The finger localization task involves lightly touching the tip of one or two fingers with a sharpened pencil point. Children are asked to indicate which finger was touched by touching that finger with the thumb of the same hand (uncrossed condition) or by touching the corresponding finger on the opposite hand with the thumb of the opposite hand (crossed condition). There are three conditions: 1) hands in view, one finger stimulated; 2) hands out of view, one finger stimulated; and 3) hands out of view, two fingers stimulated consecutively. Each of these three conditions consists of four 16-trial blocks, two uncrossed blocks followed by two crossed blocks. Starting hand is counterbalanced across subjects. Children's hands are placed in a vertical plane with palms facing each other, thus providing a mirror image of each other and minimizing directional confusion.

Vision is excluded by blindfold. The two-finger condition is included to manipulate task complexity. For this condition children are asked to "touch the same fingers, in the same order" with the thumb of the same hand or to touch the corresponding fingers on the opposite hand with the thumb of the opposite hand. For all conditions, children are told which hand was to be stimulated, how many fingers will be touched (1 or 2), and which hand they should use to respond. The examiner records responses, allowing self-correction. No feedback is given concerning accuracy of performance on trials.

Dexterity and handedness:

Grooved Pegboard Test (Matthews & Klove, 1964)

The Grooved Pegboard Test consists of a pegboard with 25 grooved slots. The child is required to place a grooved peg into each of the slots first with the right hand and then the left hand. The time to complete the task with each hand is recorded. Normative data are available for children ages five to 16. Administration time is approximately 8 minutes.

Edinburgh Handedness Inventory (Oldfield, 197)

Handedness is examined using the Edinburgh Handedness Inventory, in which each child is asked which hand they use for the following tasks: writing, drawing, throwing, using scissors, using a toothbrush, using a knife (without fork), using a spoon, using a broom (upper hand), striking a match, and opening a box (lid).

Emotional functioning and behavior:

Achenbach Scales (Achenbach, 1991).

There are three Achenbach scales that make up a comprehensive evaluation of childhood behavioral functioning, the Child Behavior Checklist completed by the parent or caregiver, the Teacher Report Form completed by the teacher, and the Youth Self Report, completed by the child (ages 11-18). These scales will be translated into Afrikanns for use in South Africa.
Child Behavior Checklist. The Child Behavior Checklist (CBCL) is a 113-item questionnaire that measures a child's competence and problem areas as perceived by the parent or caregiver [7]. There are three responses (not true, somewhat or sometimes true, very true or often true) and eight syndrome scales (Withdrawn, Somatic Complaints, Anxious/Depressed, Social Problems, Thought

Problems, Attention Problems, Delinquent Behavior, and Aggressive Behavior). In addition, for children age 6 and older, three competence scales measure the child's perceived competence in activities, social, and school settings. The eight syndrome scales are summarized in three scales: Internalizing (Withdrawn, Somatic Complaints, and Anxious/Depressed), Externalizing (Aggressive Behavior, Delinquent Behavior), and Total problem scales. The dependent measures are the Summary scales (Internalizing, Externalizing, Total) and the eight problem scales.

Teacher Report Form. The Child Behavior Checklist-Teacher Report Form (CBCL-TRF) is a companion measure to the CBCL and is completed by the child's teacher [8]. The same scales as the parent version are produced. The dependent measures are the same as for the CBCL.

Youth Self Report. The Youth Self Report (YSR) is the child version of the CBCL and is completed by children aged 11-18. It requires a 5^{th} grade reading level or can be administered orally. It has competence and problem scales parallel to the CBCL and uses the same 3-point scale as the other versions. The dependent measures are similar to the other forms and include two competence scales (Activities and Social), eight problem scales, and three summary (Internalizing, Externalizing, and Total) scores.

The Parent/Teacher Disruptive Behavior Rating Scale (Pelham, 1992)

The Parent/Teacher Disruptive Behavior Rating Scale (Pelham et al. 1992) is a parent/teacher-rated questionnaire that consists of 45 items, which are scored on a 4-point scale (Not at all, Just a little, Pretty much, and Very much). Pretty much and Very much responses to problem behaviors provide the requisite information for making DSM-IV diagnoses of attention deficit disorder, conduct disorder, oppositional defiant disorder. The test can be scored either by counting endorsed symptoms for each disorder or by comparing the target child's factor scores.

Data analysis and interpretation:

The primary objective of using a core test battery cross-culturally is to determine what tests consistently discriminate between children with FASD and controls. We will analyze the data pertaining to this specific aim in 4 steps: 1). We will conduct exploratory analyses (Tukey, 1977) with a view to checking if the assumptions of multivariate analysis of variance (MANOVA) are violated. We will also perform a series of independent group t tests in the initial stage to determine if the two groups are comparable with respect to demographic variables. 2). A 2 (site) x 2 (group) MANOVA will be performed, especially to find out if there is a site x group interaction. 3. If the interaction is not significant, then we will focus on the effect of group on a linear combination of the dependent variables. If the groups differ on specific demographic variables, those variables will be included as covariates. The merits of using the multivariate approach to analyzing the data pertaining to Specific Aim 1 are as follows: I). MANOVA guards against the inflation of type 1 error rate that may occur as a result of performing multiple univariate tests; II). MANOVA takes into account the intercorrelation between variables; and III). MANOVA offers a richer analysis of the data by revealing *systems of variables* that discriminate between the two groups. 4). Finally, if the omnibus multivariate test is significant, two follow-up procedures will be conducted to determine the relative significance of each dependent measure. Proposed by Wilkinson (1975), one procedure involves performing successive MANOVAs in which one dependent measure is left out in each analysis. The change in the multivariate F is examined to determine which variables are contributing the most to the multivariate effect. The second procedure, which is recommended by Huberty and colleagues (Huberty & Morris, 1989; Huberty & Smith, 1982), is based on the 'F-to-remove' statistic for each variable from a discriminant function analysis. This statistic reflects the decrease in each group separation if a given variable is removed from the entire set of variables.

Thus, the F-to-remove values can be rank ordered to determine the relative significance of dependent variables. The strength of these two procedures is that they take into consideration the interrelations of dependent measures. We have previously used these two procedures to determine relative contributions of dependent variables to a multivariate effect (Adnams et al, 2001).

If MANOVA revealed a group x site interaction (Step 2), one possibility would be that children in one site failed some tests because of cultural effects. This hypothesis can further be explored by isolating the source of the interaction. On the other hand, if the main effect of group is significant and if the same variables contribute to the separation of the two groups irrespective of the site, then one can make a case for prenatal exposure effects on cognitive dysfunction.

D.4 Specific Aim 2:

The second goal of the current research is to test a hierarchical model of the topography of cognitive ability in children with FASD. Hierarchical models of cognitive abilities (e.g. radex model) posit that complex tasks that are at the top of the hierarchy load on what is known as 'g' factor. We hypothesize that those complex tests at the top of the hierarchy highly discriminate between children with substantial prenatal alcohol exposure and normal controls.

Procedure:

The strategy that we plan to use to test the above hypothesis involves administering a carefully designed battery of tests consisting of 'complex' and 'simple' tests. The complexity of a test will be determined a priori based on the number of cognitive processes involved in it. The following tests have been proven to be useful in the study of the effects of test complexity.

Progressive Planning Test (Kodituwakku, 1993)

In this test the participants are required to move 3 to 5 colored beads that are arranged in a specific order in an initial position on one of three stakes to create a series of prespecified new arrangements on another stake (goal positions). The moves are subject to two rules: (1) only one bead can be moved at a time, and (2) once removed from its initial position, a bead must not be returned to that position. These rules constrain the path to the goal, increasing the working memory load in some problems. We have established from several research projects that problems 1 through 4 are simple in that they do not involve the manipulation of information in working memory. In contrast, problems 5 through 8 involve mental manipulation and hence, are complex. Two scores will be computed to reflect differential performance of children with FASD on simple and complex problems. Psychometric properties of the test are desirable (split-half reliability= .84).

Phonemic and Semantic Fluency Tests (Korman, Kirk, & Kemp, 1998)

The phonemic and semantic fluency subtests from the NEPSY: A developmental neuropsychological assessment will be used to assess verbal fluency of the participants. In the Phonemic Fluency subtest, the child is asked to generate words beginning with the letter 'S' and then those beginning with 'F' excluding names of people or places. The sum of words generated for each letter within 1-minute is taken as the phonemic fluency score. Phonemic fluency is contrasted with semantic fluency, in which the child is asked to generate exemplars of two semantic categories: animals and foods. The total number of exemplars generated for each category within a minute is summed to compute the semantic fluency score. The phonemic fluency test is more demanding than the semantic fluency test because the former requires more cognitive processes than the latter. The cognitive processes involved in phonemic fluency include retrieving words,

checking whether a word satisfies the constraints imposed by the examiner (e.g. no names of people or places), and keeping a mental record of words already produced.

Logical and spatial memory tests (Lhermitte & Sigonoret, 1972)

The spatial memory test consists of a 3 x 3 grid and nine line-drawn pictures of common objects. The object of the test is to learn the location of each drawing on the grid. The examiner places a drawing on its prespeified location and displays it for 5 seconds. Then the examiner removes it and places another drawing on its prespecified location. This procedure continues until all the drawings have been displayed. Then the examiner presents the drawings to the subject in a random order asking him or her to point to the locations. An incorrect response is immediately corrected by displaying the drawing on its correct location. The number of correct responses over 3 learning trials is the spatial memory score. In the logical memory condition, the subject is required to learn a logical pattern on the 3 x 3 grid. The logical pattern is made up of geometric shapes that vary by color and shape. In the top row of the pattern are triangles, in the middle row are circles and in the bottom row are squares. The three columns from the left to right are in blue, red, and yellow respectively. Given that this pattern is broken into 9 pieces and each piece is removed after displaying for 5 seconds, the subject has to create the pattern in the head. As in the spatial memory test, the number of shapes correctly placed over 3 learning trials is computed.
Furthermore, results from the extra and intra-dimensional set-shifting subtest of the CANTAB battery will also be used in testing the hypothesis of test complexity.
Data analysis and interpretation:

Analyzing the data pertaining to Specific Aim 2 will be conducted in two steps. The focus of the first step will be to determine if the complex tests highly discriminate between the two groups while the simple tests do not contribute to the group separation. A simple strategy to test this hypothesis would be to transform the data into z scores, to create two composite scores for complex and simple tests by computing means, and then to run a 2 (group) x 2 (tests) repeated measures MANOVA. If only complex tests discriminate between the groups, then a significant group x tests interaction can be expected.

The focus of the second step of data analysis will be on testing the hypothesis related to the topography of cognitive ability. Non-metric multidimensional methods developed by Shepard (1962) and Kruskal (1964) will be used to elucidate the hierarchical structure (radex) of cognitive abilities. Subtests from the Leiter , the Progressive Planning Test, Verbal Fluency tests, ID/ED tests from the CANTAB battery, and Logical and Spatial Memory tests will be included in non-metric multi-dimensional scaling.

The expected results of this analysis are presented in Figure Appendix A. As the figure shows, complex tests scale in the middle of radex and simple tests scale in the periphery.

D.5 Pitfalls and Alternatives:

Although obvious confounding factors (e.g. head trauma, seizure disorder, drug and alcohol abuse, and group disparities in SES, parental education, age, sex etc.) will be controlled through the study design and covariance analysis, there can be other potential sources of variability that may complicate the interpretation of data. One potential source of error concerns the cross-cultural variability in test-taking attitudes. As Nell (2000) points out, children growing up in impoverished conditions may not be “test-wise”. It is probable that children from alcoholic families are less familiar with test taking. Therefore, we plan to use extended practice to insure that the child understands

test instructions. Another potential source of error relates to test administration. Given that there will be multiple testers at each site, there is a possibility of tester effect. To minimize the tester effect, test administration will be standardized and test administrators will be provided with extensive training prior to the beginning of the study. To determine if there is a systematic variability associated with testers, an analysis of neuropsychological data will be conducted using tester as an independent factor. For the purpose of quality control, a random sample of test protocols will be re-scored. The participants will be given breaks as needed to insure their cooperation and to reduce variability of test performance.

D.6 Power analysis:

Data were available from a number of studies of cognitive function in FASD children, which were used as a basis for power analysis. Pertinent to Secific Aim 1, a study conducted in South Africa indicated that children with FAS earned lower scores on a range of tests that were comparable to those proposed in the current research. The Raven Colored Matrices Test, which contains items that are analogous to those from some subtests of the Leiter (e.g. Sequential Order, Repeated Patterns), was utilized to assess non-verbal intellectual ability. The estimated effect size measure "d" indicating the magnitude of the group effect on intellectual ability was .63. The average effect size for tests tapping different elements of attention (e.g. Digit Span, Letter Fluency) was .65. The group differences in visual-motor integration, memory, and behavior were found to be significant, as reflected by effect sizes ranging from .48 to .74. In a study conducted in the Unites States, the Wisconsin Card Sorting Test and the Progressive Planning Test were utilized to assess extra-dimensional set shifting and planning ability respectively. The estimated effect size measures for these two tests, which are comparable to the ED and the Stockings of Cambridge (SOC) subtests from the CANTAB battery respectively, were large (.83 and 1.11). Accordingly, the results from the above previous studies were used to estimate the magnitude of the multivariate effect size of the group effect on the tests proposed in the current research. In a two-group MANOVA (Specific Aim 1), Mahalanobis distance D^2 provides a natural squared generalization of the univariate measure "d" (Stevens, 1980), where individual mean scores have been replaced by a mean vector and standard deviation s has been replaced by the sample covariance matrix S. A comparison of the two groups from our wave II research in South Africa (FAS N=64; Control N=64) on a linear combination of 8 tests (non-verbal intelligence, attention, memory, and visual motor integration) revealed a highly significant group effect (Hotelling's $T^2 = 56.17$). The magnitude of the multivariate effect size measure, Mahalanobis distance D^2, is 1.92, which is very large. Therefore, the projected sample size of 200 per group would result in a power exceeding .99 for detecting the group difference at the alpha level .05. Given that the data from the current study will eventually be combined with those collected at other sites, group differences can be detected with sufficient power even with a larger set of dependent variables. A simpler strategy was employed to estimate power of detecting the differential performance of children with FASD on simple and complex tests as predicted under Specific Aim 2. First, scores from relatively complex and simple tests were transformed into z scores and then by calculating means of z scores from complex and simple tests separately two composite scores were created. An analysis of group differences in difference scores (complex –simple) indicated there was a complexity X group interaction with the effect size measure d being .48. If 200 subjects are included in a group, an effect size of this magnitude will yield a power estimate exceeding .99 at the .05 alpha level.

D.7. Timetable or sequence for proposed studies:

The proposed project will be conducted over a 5-year period. The first six months of the project will be devoted planning activities. These will include hiring staff, purchasing test materials and necessary equipment, and initiating contacts with possible additional referral sources. The efforts

will be focused on conducting training and practice sessions for research staff and recruitment of potential participants in the second half of the first year. Main activities over the next four years will be subject recruitment, eligibility screening, neurobehavioral assessment, scoring of test protocols, data entry and verification and data analysis.

D.8 Future directions:

The merits of the current research project include access to two human groups with a very high prevalence of FASD and the fact that extensive interdisciplinary studies have already been completed and are underway will help in the meaningfulness of the results. The use of a core test battery that will be used in other international sites. The use of common test battery will afford, for the first time, a rare opportunity to create a large database obtained from accurately diagnosed children with FASD, many of whom will be from the severe end of the FASD spectrum. A large dataset created with data obtained in cross-cultural settings will allow researchers test hypotheses that they were unable to test previously (e.g. hypotheses pertaining to cultural effects on test performance in FASD children and those pertaining to interactive effects of multiple factors on cognition). Despite this advantage, the multi-site data set will lack some information necessary for delineating a behavioral phenotype. For example, we did not include culture-specific tests such as verbal tests in the test battery. However, in a yet to be published study conducted in South Africa, we found that language-related abilities highly discriminated between children with FAS and controls. Specifically, we obtained evidence that children with FAS are impaired in grammar comprehension and phonological working memory. Accordingly, a systematic study of language skills will be necessary to delineate the behavioral phenotype of FASD. Educators and neuropsychologists have long recognized that how a child attains a solution to a problem is as informative as the solution itself. In other words, a response to a test item and the processes leading to the response both provide useful clues to brain organization. For example, children with Williams syndrome (WS) perform proficiently on tests assessing language and face processing. Karmillof- Smith (2000) has reported that children and adults with WS reach their behavioral mastery of face processing through processes different from those used by typically developing children. Therefore, a process analysis of test performance provides useful information for defining behavioral phenotypes. The emphasis of the test battery proposed for the current research is on behavioral end points rather than processes. We plan to incorporate process analysis of test performance in our future research on FASD.

5. Human Subjects Research

Risks to the Subjects:

Human Subjects Involvement and Characteristics:

The proposed research will involve a minimum of 400 children (FASD and controls) or a maximum of 500 children (depending on recruitment in Indian Country), 250 with FASD and 250 controls. They will range in age from 7 to 18. We expect that the South African cohorts will include relatively equal numbers of male and females (approximately 150 males and 150 females) whereas the American Indian group may comprise more predominantly males (120 males and 80 females). As noted above, the South African cohort will be made up of "coloured" children of mixed ethnic heritage. The majority of potential participants in this research project will be from lower socio-economic strata (SES), and measures of SES will be collected for each mother and child pair. Included in the FASD group will be the children

who display a continuum of dysmorphia (ranging from severe to none) and who have confirmed prenatal alcohol exposure. We will use the following exclusionary criteria in the selection of potential participants: moderate to severe mental retardation, a history of head trauma with loss of consciousness for a duration longer than 10 minutes, neurologic illness (e.g. seizure disorder), and current major psychiatric problems (e.g. psychosis), and substance abuse. The FASD and control groups will be comparable with respect to age, sex, SES, and ethnic background.

Sources of Materials:

We will collect data pertaining to cognitive abilities of the children through standardized neurobehavioral tests. Data on their emotional functioning and behavioral issues will be obtained from both teachers and parents by means of rating scales. Parents will also be interviewed to obtain information regarding children's health status and other relevant postnatal variables.

Potential Risks:

Risks associated with current research include emotional stress that be caused by test taking and loss of privacy. The data collected in this research will be handled as confidentially as possible. Case numbers will be assigned to the participants and these case numbers without any personal identifiers will be used when sharing the data with other sites. Test data will be kept in locked cabinets in secure areas. In order to reduce fatigue during testing, the participants will be given breaks as needed.

Adequacy of Protection against Risks:

Recruitment and Informed Consent:

In the South African study site, a social worker in the community or the child's classroom teacher will explain the study to the parents of potential participants because a significant number of parents have limited education. In the Native American communities, research site coordinators will contact parents/children through a letter that explains the study. The control group will be recruited through advertisements and announcements in tribal centers and community gatherings.
We will obtain informed consent from the parents/legal guardians and the assent from the children. The proposed research protocol and consent/assent forms are currently being reviewed by the Human Research Review Committees at the University of New Mexico and the Ethics Committee of the University of Cape Town. Additionally, we have already secured approval (e.g. sinlge site assurances) to conduct the study from school principals and the Regional Office of Education in Cape Town. Approval from the tribal councils of the participating American Indian sites is currently being sought. We have along term working relationship on FAS research and prevention with all six reservations, each backed by official tribal resolutions. This particular battery of tests is, however, an extension of tests already being used as part of our epidemiology studies and prevention trial at each site, and the diagnostic services that we currently provide. We therefore will not proceed until

new or amended tribal resolutions are secured. We also have existing approvals from the U.S. Indian Health Service IRB's in the Aberdeen Area (ND and SD) and Billings Area (MT) as well, and we will secure additional approvals for this particular set of neuropsychological tests.

Potential Benefits of the Proposed Research to the Subjects and Others:

There are two potential benefits to the participants. One, incentive fees ($40 US; R233 worth of food staples in South Africa) will be paid to the parents of the children (experimental and control) participating to compensate the for their time and effort. Two, each guardian will receive a letter summarizing the findings (strengths and weaknesses) of the tests for their children. In South Africa, meetings with the parents have been planned for this feedback to account for lower levels of literacy.

Importance of Knowledge to be Gained:

It is hoped, however, that the knowledge gained through this research will be useful for improving diagnosis of children with FASD, specifically those children who do not display dysmorphia. Knowledge gained trough this research will also be potentially useful for developing evidence-based interventions for the affected children.
Inclusion of Women:

It is expected that about 40 percent of potential participants are girls.

Inclusion of Minorities:

The potential participants from the United States will be American Indian children from the Northern Plains states. The participants from South Africa will be colored children from underprivileged backgrounds.

Inclusion of Children:

All participants of the current research project will be children. Two reasons necessitated the inclusion of children in the study: 1. A reliable diagnosis of FASD can only be made in children; 2. Given that adults with FASD can adversely be affected by conditions other than prenatal alcohol exposure (e.g. substance abuse, head injuries), children are better candidates for neurobehavioral studies.

6. Vertebrate animals

N/A

7. Literature Cited