A VALIDATION OF AN ETHIOPIAN VERSION OF THE LABORATORY TEMPERAMENT ASSESSMENT BATTERY

By

STEPHANIE LEA GRANT

Bachelor of Science in Psychology
Southern Nazarene University
Bethany, OK
2004

Masters of Arts in Marriage and Family Therapy
Southern Nazarene University
Bethany, OK
2006

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2008

A VALIDATION OF AN ETHIOPIAN VERSION

OF THE LABORATORY TEMPERAMENT

ASSESSMENT BATTERY

Thesis Approved:

David Thomas

Thesis Adviser

Maureen Sullivan

James Grice

Laura Hubbs-Tait

A. Gordon Emslie

Dean of the Graduate College

ACKNOWLEDGMENTS

Special thanks go to a variety of individuals who have provided a great deal of assistance on this project. First, to my advisor, Dr. David Thomas, for his guidance, patience, and proof-reading skills. I appreciate his willingness to teach. Second, to my committee members, Drs. Sullivan, Grice, and Hubbs-Tait for their instruction and time. I would also like to thank the undergraduate students who assisted on this project, Devon Eldridge, Katie Knight, Jenny Crossland, Anna Overhulser, Amanda Mayo, Sarah Whittenberg, Cesiley French, Stephanie Lee, and Julia Conder – your assistance was invaluable. Finally, to Meron Girma, Getahun Ersino, and Fikadu Reta at Hawassa University in Ethiopia for the time they spent collecting data for this project. Only with you is my work meaningful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Temperament is a well-researched construct that is often of interest when attempting to account for individual differences in data. Several models explaining temperament exist. Temperament has been defined by Rothbart (1981) as "individual differences in reactivity and self-regulation" (p. 569). Buss and Plomin (1984) have suggested that temperament is a heritable trait and a biological predisposition while Lerner and Lerner (1983) have made the case that temperament is, at least in part, environmentally determined. In addition, Goldsmith (as cited in Rothbart, 1986) has defined temperament as more of a behavioral style, which is of particular importance as Goldsmith also used this behavioral framework of temperament in his development of the instrument discussed here. Regardless of the definition used, however, studies of temperament that focus on early infancy in particular, like the current study aimed to do, remain of high interest because they are able to capture temperamental behaviors before the environment is able to play a large role in the developmental process (Matheny, Riese, & Wilson, 1985).

While temperament has found its way into an abundance of research, the assessment of temperament, particularly cross-culturally, seems to lack consistency and

reliability. Continued cross-cultural study is yet very relevant due to the increasing amount of literature suggesting the developmental importance of the construct (Rothbart, 2007). For instance, temperament has been shown to be related to factors such as empathy and conscience (Rothbart, 2007), behavioral problems, personality, neural structures and neurotransmitters, iron deficiency, and general malnutrition. The purpose of the present study was thus to provide evidence of the validity of an Ethiopian version of an infant temperament assessment using the Laboratory Assessment Battery (Lab-TAB) originally designed by Goldsmith and Rothbart (1988), as well as validity for a shortened version of this cross-cultural measure. This was done by examining the psychometric properties of the instrument with a sample of 6-month old Ethiopian infants.

CHAPTER II

REVIEW OF LITERATURE

Temperament Assessment

Traditional temperament dimensions that are assessed psychometrically can include negative behaviors such as fear, distress to limitations, or overall negative emotionality; positive behaviors such as smiling, laughter, or duration of looking; and more neutral behaviors such as activity level, threshold, intensity, rhythmicity, approach, adaptability, attention span, distractibility, or persistence. Overlaps among some of these dimensions are present, such as that noted between fear and attention (Rothbart, 1981; 2007). However, these dimensions do appear to be independent of each other (Lemery, Goldsmith, Klinnert, & Mrazek, 1999).

Stability of temperament. There have been a few consistently raised concerns regarding temperament assessment. One of these is related to the stability of temperament over time. Temperament has been reported to be generally stable, though stability varies within factors, with some aspects of temperament seemingly more stable than others (Lemery et al., 1999; Pedlow, Sanson, Prior, & Oberklaid, 1993). Most factors do appear to increase in stability over time (Pettit & Bates, 1984), and while infancy is thought to be the least stable time period in this regard (Lemery et al., 1999), some aspects of temperament, such as activity level, demonstrate quite high levels of stability

during this period (Buss & Plomin, 1975). Irritability and emotionality also appear to be fairly stable during infancy (Riese, 1987), with objective measures such as smiling, looking, and vocalizing showing lower levels of stability than subjective measures (Pettit & Bates, 1984). Parental report of temperament is also consistently reported to be stable, though this finding leads more to questions of the validity of these assessments that rely on such information than confidence in the actual stability of the constructs (Lemery et al., 1999; Pettit & Bates, 1984).

Reliability and validity of temperament. Another concern within the area of temperament research relates to the reliability and validity of various temperament assessments. There is no "standard" temperament assessment. Instead there are a variety of widely used assessments that are specific to age groups. This in itself makes longitudinal studies difficult because not only are assessments changing over time, but there is not a lot of evidence that each measure assesses the same variables (Lemery et al., 1999). This also poses a unique problem in selecting a measurement because there are not only various assessments to choose from, but different methodologies of assessments are available as well, such as parental reports and laboratory assessments.

Parental assessment

Parental assessment has been one of the most widely used temperamental assessment measures because of its ease of use and cost effectiveness, most often relying on maternal report. Examples of such assessments include the Infant Behavior Questionnaire (IBQ, Rothbart, 1981); the Infant Temperament Questionnaire (ITQ, Carey, 1970), the Revised Infant Temperament Questionnaire (RITQ, Carey & McDevitt, 1978; Rothbart, 1981), or the Shortened Version of the RITQ (SITQ, Sanson, Prior,

Garino, & Oberklaid, 1987); the Infant Characteristic Questionnaire, which is a measurement designed specifically for assessing difficult temperament (ICQ, Bates, 1989); the Early Infancy Questionnaire (EIQ, Medoff-Cooper, Carey, & McDevitt, 1993); and the Baby Behavior Questionnaire (BBQ, Bohlin, Hagekull, & Lindhagen, 1981).

Hubert, Wachs, Peters-Martin, and Gandour (1982), and more recently Slabach, Morrow, and Wachs (1991) have conducted thorough reviews on the psychometric properties of various parental assessment measures. Moderate internal consistency appears to be the norm for most infant temperament measures (Hubert et al., 1982; Slabach et al., 1991), with the IBQ demonstrating the highest level of internal consistency. The RITQ shows the most variability of internal consistency, with the approach domain most often reported as achieving the highest levels and sensory threshold, adaptability, and intensity the lowest levels. Satisfactory test-retest reliability has been found for both the RITQ and SITQ, with the full version of the RITQ demonstrating satisfactory levels even across cultures.

Despite the widespread use of maternal report, serious questions have been raised regarding the validity of relying on such data, as was noted briefly above. While Lemery (1999) holds that maternal reports are valid, mother- v. observer-based ratings have been shown to typically possess low convergent validity (Bornstein, Gaughran, & Segui, 1991; Hagekull, Bohlin, & Lindhagen, 1984) and parents appear to have formed stable expectations of their children prior to birth, making valid assessments difficult (Mebert, 1991) and the results subjective with both perception and objective measurements.

Laboratory Assessment

An alternative method of temperament assessment is laboratory assessment. Laboratory procedures have the advantage of greater experimental control and objectivity, but may elicit uncharacteristic behaviors from infants due to the unfamiliar setting and stimuli (Bornstein et al., 1991). Goldsmith and Rothbart (1988) created an instrument for assessing early temperament in a standardized format within a laboratory, which is currently the only standardized laboratory assessment available. The Laboratory Temperament Assessment Battery (Lab-TAB) assesses joy/pleasure, fearfulness, interest/persistence, anger proneness, and activity level within children, viewing temperament across dimensions, rather than by typology. Currently, three versions of Lab-TAB exist. Two versions have been designed for assessing temperament in infancy, a prelocomotor and a locomotor version, and one has been designed for assessing temperament in preschoolers. The ideas behind the infant versions were originally conceived by Goldsmith and Rieser-Danner (University of Wisconsin Twin Center (UWTC), n.d.). Goldsmith and Rothbart then later formalized the design and coding process to what it is currently. Goldsmith and Reilly later designed the preschool version. (See the UWTC website for full versions of these assessments.)

Lab-TAB is comprised of twenty, 3 to 5 minute episodes designed to elicit temperament behaviors that would be found in everyday situations (UWTC, n.d.). Such behaviors include those related to emotional expressivity, approach/avoidance, activity level, and self-regulation. Facial, vocalic, and motoric behaviors are all rated as indicators of targeted emotions and are coded numerically. Variables are coded using different levels of measurement. Some variables are coded categorically as either present or absent

(e.g. presence of positive vocalization). Some variables are coded on a Likert scale to assess peak intensity (e.g. intensity of smiling). Other variables are coded on a ratio scale to measure time until the occurrence of a behavior (e.g. latency to fear). If time elapsed while the infant was engaging in a behavior is coded, it is converted to an interval level of measurement by way of a Likert scale (e.g. duration of looking).

Within each of the 20 episodes are various numbers of epochs by which the episode is broken down into smaller time segments for coding purposes. If an episode is broken down into multiple epochs, the same behaviors are being coded uniformly each time, though different things may be occurring in different trials (e.g. toy on the table v. toy in the experimenter's hand). Within some episodes, the coding is broken down first into intervals, trials, or stages and then is further broken down into epochs, again for ease of coding. Here, an interval represents a passage of time, a trial can refer to either a passage of time or to the fact that the experimenter is behaving differently in the episode (e.g. standing v. sitting), and a stage represents an escalation in what is occurring within the episode (e.g. stranger moves closer to the infant in progressing stages). If an episode is broken down into intervals, trials, or stages, epochs within each are averaged together such that an episode results in as many averages for a coded variable as there were intervals/trials/stages within that episode. If an episode is not broken down into anything other than epochs, the epochs of a variable are averaged together for one result. Only averaged data are used for analysis purposes.

Though the episodes are designed to elicit a specific emotion, they often reliably elicit additional emotions as well. An example of this would be an episode designed to elicit anger that often provokes sadness. As such, in some of the episodes, these

additional emotions have been included for scoring. In addition, other variables of interest are provided for coding within each episode (i.e. parent behavior, and baseline state). These variables do not address the temperament domain assessed, but are often necessary for consideration when working with infant data.

One purpose for the development of Lab-TAB was to negate the need of creating original laboratory temperament measures for each new research study (UWTC, n.d.). The design of Lab-TAB allows individual researchers to modify the situations, as well as the coding system, to best fit their current studies. Coding of the episodes yields data that vary across levels of measurement (i.e. present or absent, peak intensity, latency), also allowing for individual tailoring. In addition, by utilizing a shared measurement, standardization can be conducted on larger samples with the hope that this would eventually result in more accurate scores and allow for more readily comparable results across studies.

Psychometric properties. A drawback of allowing the customizability of Lab-TAB is related to the difficulty with standardization, however. Various studies have used modified versions of Lab-TAB (e.g. Bridges, Palmer, Morales, & Hurtado, 1993; Buss & Goldsmith, 1998; Goldsmith & Rieser-Danner, 1990; 1991; Hane, Fox, Cindy, Ghera, & Guner, 2006; Kochanska, Coy, Tjebkes, & Husarek, 1998; Rothbart, Ahadi, & Evans, 2000), many of them reporting individual psychometric properties. As Lab-TAB was adjusted for the purposes of each researcher's study, however, the psychometric properties may likely not extend to studies customizing the protocol in a different manner. However, Goldsmith and Rothbart (1991) report interrater agreement for Lab-TAB that ranged from .87 to 1.00. Bridges and colleagues (Bridges et al., 1993)

examined Lab-TAB and the IBQ for convergent validity and found that the data did provide convergent, but nonidentical, information about infant temperament. Two examples of this are that observed anger correlated with reported distress to limitations, and pleasure expressions correlated to reported smiling and laughter. Also noteworthy from their study was that assessments of positive or negative temperament aspects were often related to the more neutral aspects of temperament, such as duration of looking. Bridges et al. note that it is not surprising that infants who were prone to attend to an object for a lengthier period of time were also those able to tolerate higher levels of frustration and who exhibited higher levels of positive affect.

Hane and colleagues (2006) examined the moderating role that observed infant affect during maternal interactions had on maternal and laboratory temperament ratings using the IBQ (Rothbart, 1981) and the Locomotor version of Lab-TAB (Goldsmith & Rothbart, 1999). With temperament variables related to infant negativity, ratings converged between maternal and observer ratings when infants were involved in routine home-based activities and when negative affect was high. Ratings of positive affect converged when such affect during play was low. This finding suggests that negative behaviors either may be more easily agreed upon or more consistent across situations than positive behaviors.

With infant temperament still a necessary topic of interest, Lab-TAB appears to provide a suitable alternative for cross-cultural research to the many parental response measures of infant temperament that are currently under critique for their validity and which pose difficulties with translations. With the literature on cross-cultural

temperament research scarce, it is thought that Lab-TAB might be able to provide an appropriate direction for continuing this line of study, as will be further discussed below.

Cross-Cultural Temperament Research

Lab-TAB has still not been widely used, and no known studies have used Lab-TAB in cross-cultural research, though there has been research conducted cross-culturally with other assessments (e.g. DeVries & Sameroff, 1984; Gasman et al., 2002). The one exception to this is the Kennedy et al. (2008) and Shaw, Grant, and Thomas (2008) work in Ethiopia that is the basis for the current study and which will be discussed in more detail below. Cross-cultural research in this area has unique difficulties because not only is there a great deal of variability in the treatment of infants across cultures (Rothbart et al., 2000), but translations of traditional parental report measures are difficult, in part due to the varying cultural expectations and assumptions of infant behavior and development.

One study that attempted such a translation of a parental report measure was conducted by Boer and Westenberg (1994) using the Emotionality, Activity, and Sociability Questionnaire (EAS). In a sample of Dutch children of ages 4 to 13 years old, Boer and Westenberg found the EAS to be valid for use in this population. Mathiesen and Tambs (2000) also used the EAS in a Norwegian sample of children aged 18, 30, and 50 months, finding that temperament predictability and gender differences both increased with age.

A cross-cultural study of particular importance to the current research was conducted by deVries and Sameroff (1984) on infant temperament in three east African societies of Kenya. While still more in-depth descriptions of the societies are given by

deVries and Sameroff (pp. 84-86), rather lengthy summaries will be given here due to their significance and applicability to the Ethiopian culture examined in the present study.

The first tribe, the Bantu Kikuyu, located west of Nairobi, Kenya, were employed as wage laborers and small farmers and considered the most modern of the three groups (DeVries & Sameroff, 1984, p. 84). "Nutrition and general health [were] relatively good" within this society where housing ranged from the more traditional, round dwellings with thatched roofs to apartments (p. 84). Infant care was shared by the mother with other Kikuyu individuals, and child-rearing relationships varied from monomatric to polymatric, with men absent from home life most of the day. Infants were viewed as vulnerable, and weaning occurred earlier than in other Kenyan societies and training interactions (e.g. motor or grooming skills) occurred later, with low expectations of social and motor achievement.

The second tribe, the Bantu Digo, located south of Mombasa, Kenya, lived in "extended family clusters" and worked as farmers or fishermen (DeVries & Sameroff, 1984, pp. 84-85). Their houses were primarily large rectangular dirt structures with grass roofs where 6 to 10 family members lived together. These homes were typically organized into villages with life largely sexually segregated and men absent most of the day, as it was with the Kikuyu. The first two months of a Digo infant's life was spent in "almost constant physical contact with the mother" (p. 85). At around 2 months of age, young siblings and other individuals "gradually increase[d] their caretaking responsibilities" (p. 85). "The Digo [felt] that the infant [was] ready to learn at birth and expect[ed] a high degree of motor and social achievement at 3 to 5 months" of age (p.

85). Mothers also felt "that responding to a newborn's cry or behavior [was] a method of instruction," and thus "monitor[ed] infant behavior closely" (p. 85).

The third tribe, the Masai, lived south of Nairobi, Kenya and into Tanzania (DeVries & Sameroff, 1984, pp. 85-86). They were a pastoral Nilohamitic people who worked as subsistence herders, thus centering life around cattle. The Masais had been affected little by modernization and lived as strict traditionalists in mud-wattle structures. "Family groupings…ranged from brothers or friends living together to groups of mothers, warriors, and children" (p. 85). Infectious disease rates were high, though infants were not viewed as particularly vulnerable. Demand breastfeeding of infants continued for 2 to 3 years and formal training was not focused on, with the Masai feeling infants learned through observation of others. The society was patrilineal with a warrior emphasis, with "aggression and assertiveness…encouraged in young boys" (p. 85).

de Vries and Sameroff (1984) used the ITQ after appropriate translations had been made for each tribe and looked for differences in infant temperament with respect to modernization, experience of life-events, gender, and culture. Higher levels of modernization were "associated with greater activity, more intensity, and lower thresholds" (p. 91) within the Kikuyu and Digo. When examining early life events, different temperament factors were correlated in each culture. For the Masai, "poor family health and problematic pregnancies were associated with less adaptable infants" and within the Kikuyu, "problems in pregnancy…correlated with the negativity and intensity factor" (p. 91). Though Kikuyu infants scored as more difficult, Digo women were more likely to rate their infants and pregnancies as more difficult than the mothers within the other tribes were. While no main effects were found for sex on any of the

temperament dimensions, significant interactions were found between sex and tribe, with the Kikuyu showing no sex differences, the males as more difficult among the Digo (the matrilineal culture), and the females as more difficult among the Masai (the patrilineal culture). Significant differences were found among tribes on every temperament dimension except threshold to stimulation. These cultural differences had more of an effect on infant temperament than the modernization differences among tribes did.

de Vries and Sameroff (1984) concluded that cultural affiliation was the strongest predictor of infant temperament. Some of the temperament dimensions appeared to be more sensitive to environmental stimuli. For example, negativity seemed related to "infants' reactions to customary child-rearing practices" (p. 92) and intensity more to socioeconomic influences. de Vries and Sameroff also found that the relations between life events and temperament depended in part on the environmental characteristics, such as child-rearing orientation, and the mothers' perception of infant troublesomeness was related more to cultural expectations than to the temperament characteristics assessed by the ITQ.

This study by deVries and Sameroff (1984) suggests several notable items that are relevant to the present research. First, given the variations in reported troublesomeness among tribes, reliance on parental report of infant temperament will not likely be as valid as a laboratory assessment – particularly if cross-cultural comparisons are desired. Second, culture should be taken into consideration when assessing temperament, but should also be considered for its effects on shaping temperament, even at this early age of infancy. And third, an important question is raised regarding whether culture then

modifies temperament or whether infants that are a better "fit" within the culture are more successful (Thomas & Chess, 1977).

Cross-cultural laboratory –assessment. Laboratory assessments do assist in overcoming the difficulty of translation; however, knowing what types of behaviors will elicit desired temperament traits in other cultures remains a challenge, and little work has been done using this methodology. In a study that attempted to examine the effects of malnutrition on temperament using the prelocomotor and locomotor versions of Lab-TAB (Kennedy et al., 2008; Shaw et al., 2008), attempts were made to revise Lab-TAB to an assessment measure that would validly assess infant temperament at 6 and 9 months of age in Ethiopian infants from rural communities. No known study had attempted to take Lab-TAB into any African country until this point.

To begin altering Lab-TAB for this study (Ethiopian Lab-TAB; Kennedy et al., 2008; Shaw et al., 2008), two episodes from each of the five Lab-TAB dimensions (fear, joy, interest, activity level, and anger) were chosen for use in pilot testing: 1) A stranger approached the infant and a parasol was opened in front of the infant's face to elicit fear; 2) an experimenter tickled the infant with a puppet and the mother and infant engaged in a modified peek-a-boo game to elicit joy/pleasure; 3) the infant was encouraged to engage with blocks and a female adult approached the infant to elicit interest/persistence; 4) the infant was placed in both a prone and supine position as well as being given a basket of toys to play with to assess activity level; and 5) the infant was given a toy that was then retracted from him or her and restrained in a car seat to elicit anger/frustration, though the car seat was replaced with an Ethiopian shopping basket due to the infants' unfamiliarity with a car seat. This restraint episode in the shopping basket was later

replaced by an episode conducted while the infant was restrained while being measured for length. Some of the other episodes traditionally available for inclusion in Lab-Tab were excluded from possible consideration because they involved stimuli that were expected to be extremely foreign to both the infants and mothers of rural Ethiopia (e.g. the presentation of an unpredictable mechanical toy or several slides on a video screen).

Cultural validity was initially assessed through a pilot study on 14 infants (6-12 months of age) from rural Ethiopia using nine of the 10 above-mentioned episodes (Kennedy et al., 2008; Shaw et al., 2008). Data were coded by an Ethiopian faculty member from Hawassa University and an American undergraduate student. The scoring system used assessed the extent to which each episode exhibited construct validity; i.e., was congruent with the emotion it was intended to elicit. From this pilot study, 6 of the 10 episodes were modified to improve their cultural validity. Specifically, modifications were made because several infants exhibited a degree of distress that was attributed to separation from the mother in the unfamiliar testing situation. Thus, most episodes were subsequently carried out with the infant in the mother's lap, which is unlike the standard Lab-TAB protocol in which the infant sits in a high chair in most episodes. This, however, was understandable as in the Sidama region of southern Ethiopia, infants are rarely out of contact with their mothers and high chairs are not used (see Table 1 for a summary of Ethiopian Lab-TAB protocol used in this study).

After this initial piloting trial, Ethiopian Lab-TAB data were gathered on 106 infants at both 6 and 9 months (Kennedy et al., 2008; Shaw et al., 2008) using the developed protocol. Researchers in Ethiopia were trained as the primary coding team and a secondary team in the United States was also trained to code data. To determine

interrater reliabilities, the median intraclass correlation coefficients were found on the averaged scores for each variable within an episode. Median values were used because the data were not normally distributed. When multiple trials, epochs, or stages were included within an episode, each of these averages was treated as a separate value when calculating the median intraclass correlation coefficient. Then, the resulting medians from the two episodes measuring the same temperament variable (joy, anger, fear, interest, and activity level) were calculated. The resulting interrater reliabilities for 20 randomly selected 6-month-old infants were .68, .52, .43, .85, and .55, for joy, anger, fear, interest, and activity episodes respectively. For 9-month-old infants the values were .57, .66, .43, .78, .75 (see Table 2 for individual episode interrater reliabilities).

Despite attempts with the pilot, during the coding process of this data, questions were raised about whether the individual Ethiopian Lab-TAB episodes were assessing the temperament aspects they were designed to elicit. For instance, the Stranger Approach (Episode 8), designed to measure fear, appeared often to the researchers to elicit instead a pleasure/interest response. If such differences were in fact actually present in this data, and not simply illusory correlations, they still could be considered a result of the nutritional deficiencies the study by Kennedy et al. (2008) was attempting to examine. However, they could also be explained by the lack of validity this version of Lab-TAB had after being tailored for this Ethiopian sample.

In addition, because future studies had already planned on using this version of the Ethiopian Lab-TAB again with 900 additional Ethiopian infants in large zinc supplementation study, it became apparent that more conclusive evidence was needed regarding the measure's validity following the revisions made from the original version.

The Ethiopian culture. Drawing on the work of deVries and Sameroff (1984), the researcher conjectured that there may be specific cultural differences that could account for the differences in assessment researchers reported seeing. Aboud and Alemu (1995) report that outside of the Digo tribe, social interaction among Ethiopian mother-infant dyads appeared to be lower than it is in other countries, noting a specific difference in verbal interactions in particular. They describe the mother's inability to coordinate household tasks, such as picking up and cooking, while responding actively to their infants. During infancy, mothers are the primary source of social interaction, so although a high percentage of infants came into contact with another person during the observation procedures, Aboud and Alemu note that the infants receive little contact from these outside sources.

Aboud and Alemu (1995) also found that even after controlling for age differences, higher rates of carrying and holding that were not made in response to the infants' behaviors, as was found in the Digo tribe, negatively predicted infants' performance development. Though mothers seemed to feel they were providing protection or comfort, it appears as if these lifting and holding actions actually interfered with infant development as measured by the Bayley Scales of Infant Development (Bayley, 1993), specifically as related to visual-motor coordination and verbal comprehension and production. It is possible that these actions might have served to modify temperament as well. However, child development was positively predicted by mothers' speech responsiveness which is not a behavior commonly seen within the Sidama region of Ethiopia (Getenesh Berhanu, personal communication, July 17, 2007).

Another factor that needs to be taken into consideration when examining differences in Ethiopian infants is maternal knowledge of infant development. Aboud and Alemu (1995) found that overall, those mothers who expected later development in comparison with the other mothers in the study were less verbally responsive to their infants. In reference to language development, 21% of mothers did not think children could begin to understand spoken words until 18 months of age or later and 20.5% thought they could not influence their child's ability to learn new words until 48 months of age or later. Child-rearing in Ethiopia is based on traditional values where children learn skills not by playing, but by watching their mothers. Though there were differences between tribes, Aboud and Alemu note that children are not expected to be independent and inquisitive, but instead to be non-demanding and to play quietly close to their mothers. These expectations of infants are similar to those found in the Sidama region of Ethiopia. These findings might suggest a cultural advantage to infants displaying low activity levels, high soothability, and low reactivity.

Current Study

Based on the information provided by Aboud and Alemu (1995) and the observations researchers made in Kennedy et al. (2008) and Shaw et al. (2008), it is possible that the Ethiopian Lab-TAB may not have accurately assessed temperament in rural Ethiopian infants. Possibilities for these difficulties could be that the stranger approach did not elicit fear because these infants have had different social experiences than American infants and come in contact with fewer strangers; peek-a-boo and the puppet game may not have elicited joy because either the infant was not accustomed to socially engaging with adults or mothers were not close enough to provide a sense of

security; the toy retraction and restraint episodes might not have been able to elicit anger/frustration because Ethiopian infants are encouraged to be quiet when around adults; the task orientation with blocks and person interest might not elicit interest/persistence responses because the infants were not encouraged to be exploratory or to engage with adults; and the prone and supine placement and basket of toys designed to elicit activity level may not be effective because the infants were both unfamiliar with the toy stimuli and were not expected to be inquisitive with new objects.

The primary purpose of the present study was thus to examine the validity of this new Ethiopian Lab-TAB. A secondary purpose of this study was to assess whether Ethiopian Lab-TAB could be shortened for future use. In the original study (Kennedy et al., 2008; Shaw et al., 2008), each infant assessment took approximately 30 minutes to videotape and approximately 135 person minutes for coding. If a briefer measure could be designed, it would enable future research to save on both time and money.

CHAPTER III

METHODOLOGY

Participants

Previously recorded videos of Ethiopian infants were used for the present study. No additional data were gathered. Videos were collected in a rural area of Ethiopia known as the Sidama region and then sent to the United States, where they were analyzed for the purposes presented here. One-hundred eight predominately breastfed Ethiopian infants (males = 49, females = 59) participated originally in a study examining the effects of iron deficiency on anthropometry and temperament. Infants were tested at both 6 and 9 months of age. The mean age for the 6-month-old infants was 23.16 weeks (*SD* = 2.05). Using standards adopted by the World Health Organization (2008), 22% of these infants were classified as anemic and 25% were considered stunted. Ninety-four of these infants were included in the present analyses. Fourteen infants were excluded either because the video files were corrupted and could not be viewed or the Ethiopian Lab-TAB procedure had to be stopped because the infant became upset during testing. A pair of male twins who had been born prematurely were also excluded.

Materials

Generalized Temperament Assessment. The Generalized Temperament Assessment (GTA) was created for purposes of this study (see Figure 1). Its purposes were 1) to provide a brief measure of temperament that assessed overall impressions of what the infant was displaying rather than focusing on specific behavioral indicators, and 2) to allow for the coding of any display of temperament rather than only a specific domain. This is unlike the usual Lab-TAB coding procedure (UWTC, n.d.) in which each episode was designed to allow coding for only one of the five temperament domains. The five temperament domains assessed in the Ethiopian Lab-TAB (joy, anger, fear, interest, and activity level) were used as the possible temperament domains in the GTA.

An overall subjective impression of each of the temperament domains is coded for each epoch or time period using a Likert-type scale (0 = Clearly Not Present, 1 = A Little Present/Ambiguous, 2 = Somewhat Present, 3 = Very Noticeably Present). These time periods can be modified for purposes of the researcher. In the present study, the epoch times from the original Ethiopian Lab-TAB data (Kennedy et al., 2008; Shaw et al., 2008) were used.

Shortened Version of the Ethiopian Laboratory Temperament Assessment Battery. An initial analysis was conducted to assess the feasibility of creating a briefer version of the Ethiopian Lab-TAB (Ethiopian Lab-TAB-S). Cronbach's alphas were run on the data collected in the original study (Kennedy et al., 2008; Shaw et al., 2008) with the averaged data for each variable within each of the episodes. Exceptions to this were the Parasol Opening (Episode 2) and the Stranger Approach (Episode 8). Due to the already brief nature of the Parasol Opening, no attempt was made to shorten it further. The Stranger

Approach has varying situational behaviors in each epoch that have been designed to elicit varying temperamental responses, and so it was decided it should not be shortened. It was also determined that no attempt would be made to shorten the Person Approach (Episode 7) because the experimenter was to act in different ways during the episode (talk, stand silently, or sit in a chair). Within each of the 10 episodes, calculated Cronbach alpha values for each variable coded were to be greater than .70 to be considered appropriate for shortening the episode.

Results showed that each variable coded within each of the episodes achieved Cronbach alphas of at least .70 (see Table 3). As all episodes (other than 2, 7, and 8) were deemed acceptable to shorten from this Pilot Study, shortening was done by dropping the number of epochs or trials coded in each episode. The Spearman-Brown Prophecy Formula was used to calculate an acceptable number of epochs to cut from each episode while still maintaining internal consistency values of .70 or higher (see Table 4). This formula attempts to predict the reliability of a test after changing its length (Cronbach, 1984). A chosen increase factor is multiplied by the desired level of reliability (e.g. .70). This product is then divided by a denominator of [1 + (the chosen increase factor – 1) * (desired reliability)] to calculate an estimated reliability. While use of this formula suggested that some episodes could be shortened by more than 50%, no more than 50% of an episode was cut from the protocol.

The resulting version of the Ethiopian Lab-TAB-S still consisted of 10 episodes and the same variables were coded within each of them. However, the Puppet Game (Episode 1) was shortened from 5 to 3 epochs, the Peek-A-Boo Game (Episode 3) from 6 to 3 epochs, the Task Orientation (Episode 4) from 18 to 9 epochs, the Toy Retraction

(Episode 5) from 9 to 6 epochs (3 to 2 trials), the Basket of Toys (Episode 6) from 9 to 6 epochs, the Prone and Supine Placement (Episode 9) from 6 to 4 epochs, the Restraint (Episode 10) from 6 to 3 epochs, and the Parasol Opening, Person Approach, and Stranger Approach were not shortened (Episodes 2, 7 and 8, respectively). The time the infant was being measured dropped from 15 minutes to 10 minutes with these changes. (See Figure 2 for the manual of the Ethiopian Lab-TAB-S.)

Procedure

Study 1. The purpose of Study 1 was to examine the validity of the Ethiopian version of Lab-TAB. Of interest was whether the 10 individual episodes measured the temperament domain for which they were designed. The GTA developed for this study was used for this purpose.

Following the completion of a pilot study with the GTA, two research assistants who were unfamiliar with the Ethiopian Lab-TAB individually reassessed infant temperament with the GTA by viewing the videos of 6-month-old infants that were collected in Ethiopia in 2006. All of the 6-month-old infant videos were thus reassessed two times (N=94). These assistants were unfamiliar with the Ethiopian Lab-TAB and the purposes of the original Kennedy et al. (2008) study. The time codes used in the original Ethiopian Lab-TAB assessment were inserted into the GTA coding sheets for each infant so that each Ethiopian Lab-TAB episode was re-rated by researchers on each of the five temperament domains.

Study 2. The purpose of Study 2 was to examine the reliability of the Ethiopian Lab-TAB-S when compared to the original Ethiopian version used in Kennedy et al., (2008) and Shaw et al. (2008). For this study, two teams, each comprised of two research

assistants, were trained to code infant temperament using the shortened version of the Ethiopian Lab-TAB. These research assistants were unfamiliar with the original version used in the Kennedy et al. (2008) study. They reassessed infant temperament in pairs using a conferencing method, as was originally done. The first team reassessed all of the 6-month-old infants by viewing shortened versions of the videos that were collected in Ethiopia in 2006/2007. The second coding team reassessed 20 of the videos for purposes of calculating interrater reliability.

CHAPTER IV

FINDINGS

Study 1

The purpose of Study 1 was to examine the validity of the Ethiopian version of Lab-TAB. Of interest was whether the 10 individual episodes measured the temperament domain for which they were designed using the GTA that was developed for this purpose. First, scores across epochs of the GTA were averaged for each episode of the data. Interrater reliability was then calculated using intraclass correlations for each averaged variable (see Table 5). The median correlation coefficient was then found for each episode (see Table 6). Median values were used because the data were not normally distributed. The median correlation coefficients for each domain of the GTA were .61, .72, .71, .69, and .35 for joy, anger, fear, interest, and activity level, respectively.

Descriptive statistics for the GTA results are reported in Table 7. Separate one-way ANOVAs were run for each episode measured by the GTA at alpha levels of .05 to examine differences in average scores within temperament domains (joy, anger, fear, interest, activity level; see Table 8). Note that differences in degrees of freedom are a result of missing data for that particular episode. Significant differences in scores were found for all episodes.

When examining mean scores, these analyses allow one to see that different temperament categories are differentially represented in each episode. Except for Episode 9, interest was rated highest for each episode, followed by activity level. Both of these domains are classified as neutral variables. For Episode 9, an episode designed to measure activity level, activity level was rated highest, followed by interest. For each of the six affect variables (episodes 1, 2, 3, 5, 8, and 10), the third highest rated temperament domain (after interest and activity level) was the temperament domain the episode was designed to measure.

Follow-up post hoc analyses were run using Tukey's HSD (see Table 10). When specifically examining the affect episodes and looking for differences between the affect domains, most comparisons were statistically significant. However, in Episode 3, joy and anger were not statistically different from each other; in Episode 8, fear and joy were not statistically different; and in Episode 10, anger and fear were not statistically different.

Study 2

The purpose of Study 2 was to examine the reliability of the Ethiopian Lab-TAB-S when compared to the original Ethiopian version used in Kennedy et al. (2008). As in the Ethiopian Lab-TAB data, scores across epochs were first averaged for each episode of the data from the Ethiopian Lab-TAB-S. Interrater reliability was then calculated using intraclass correlations for each averaged variable (see Table 10). The median correlation coefficient was then found for each episode (see Table 11). Median values were used because the data were not normally distributed. Then, the resulting medians from the two episodes measuring the same temperament variable (joy, anger, fear, interest, and activity level) were calculated. The resulting interrater reliabilities for 20 randomly selected 6-

month-old infants were .51, .62, .21, .78, and .73, for joy, anger, fear, interest, and activity episodes respectively.

Pearson *r*s were used to examine relations between temperament averages among episodes from the original Ethiopian Lab-TAB data and the data from the Ethiopian Lab-TAB-S at alpha levels of .05 (see Table 12). All variables were statistically significant except for one variable in Episode 2 measuring escape behaviors and variables measuring facial or bodily sadness. The median correlation coefficient for each episode was then found (see Table 13). The resulting medians from the two episodes measuring the same temperament variable (joy, anger, fear, interest, and activity level) were calculated. The resulting correlation coefficients were .66, .72, .56, .73, and .85 for joy, anger, fear, interest, and activity episodes respectively.

CHAPTER V

CONCLUSION

The purpose of the present study was to provide evidence of the validity of the Ethiopian Lab-TAB, as well as validity for a shortened version of this cross-cultural measure of temperament. Under the behavioral definition of temperament proposed by Goldsmith (as cited in Rothbart, 1986), results from the present study suggest continued validation work of Lab-TAB with an Ethiopian population as well as support for the shortened version of this measure.

Study 1

Within Study 1, significant differences between temperament domains were found for each episode. Except for Episode 9, interest was rated highest for each episode, followed by activity level. For Episode 9, an episode designed to measure activity level, activity level was rated highest, followed by interest. For each of the six affect variables, the third highest rated temperament domain was the domain the episode was designed to measure. These results suggest that the GTA did provide support that each of the 10 Ethiopian Lab-TAB episodes was eliciting the affect temperament domain it was designed to measure, which was the primary concern originally.

However, the neutral temperament domains were rated consistently highest among the other domains. One possible explanation is found in research by Rothbart (1981; 2007). Rothbart suggested overlaps among some of the dimensions were to be expected, particularly for variables measuring affect and those measuring attention or activity level, as was found in the present study. It would make sense that those episodes that elicit the strongest affect also command the most interest and activity. On the other hand, interest and activity can occur in the absence of affect. In such cases, a lack of affect cannot be attributed to a lack of engagement with the stimuli. Bridges et al. (1993) also demonstrated that assessments of emotionality were often related to the more neutral aspects of temperament, such as activity level and attention. Therefore, the higher means from the neutral variables that appeared within episodes designed to measure affect are understandable from several perspectives.

A second explanation for the high ratings of the neutral temperament domains could be that the Ethiopian infants assessed in the present study expressed flattened affect, making the affective behaviors that were present more difficult to notice, and thus more difficult to code. This could be a result of the cultural expectations for Ethiopian infants. Aboud and Alemu (1995) suggest that infants are encouraged to be non-demanding and quiet, which would likely discourage emotionality. It could also be a result of the malnutrition or iron deficiency that Kennedy et al. (2008) hypothesized might have an effect on infant temperament development. Regardless of the explanation, the idea of flattened affect is supported by the low variability and originally coded rates of most of the affective variables in the Kennedy et al. study, several of which were not present to code at all (e.g. vocalizations, laughter, smiling).

Differences in mean scores could also be related to the variations in stability of the temperament constructs during infancy. Buss and Plomin (1975) found that activity level demonstrated high levels of stability and Riese (1987) showed that general affect variables were fairly stable. However, Petit and Bates (1984) also found that objective measures such as smiling, looking, and vocalizing showed lower levels of stability than subjective measures. This might suggest that measuring temperament through a subjective assessment such as the GTA might result in higher levels of stability than measuring through standard Lab-TAB coding.

That being noted, it was the data from the GTA analyses that was measured. However, the GTA data were gathered from videos that followed the Lab-TAB protocol. One possible explanation for not seeing more pronounced differences between mean scores is that the laboratory assessment may have elicited uncharacteristic behaviors from infants due to the unfamiliar setting and stimuli associated with the tasks (Bornstein et al., 1991). Efforts were made through use of a pilot study to ensure that infants were responding appropriately to the stimuli (e.g. infants were placed in their mothers' laps and Caucasian experimenters were not utilized). However, the entire protocol was largely out of the ordinary for the infants as they are not traditionally exposed to toys or engagement with adults (Abound & Alemu, 1995).

Therefore, while questions were raised about whether the individual Ethiopian Lab-TAB episodes were assessing the temperament aspects they were designed to elicit, use of this measurement to assess joy, anger, fear, interest, and activity level is thus supported by the present study's analyses with the GTA. It should be expected, however, that affect variables will likely be low in intensity and neutral variables rated as higher

for most episodes. This might account for the concerns noted by Kennedy et al. (2008) about the validity of the measure. The specific instance noted earlier with the Stranger Approach (Episode 8) was that it was designed to measure fear but appeared often to the researchers to elicit instead behaviors of pleasure or interest. The present work does validate this concern as interest and activity level were behaviors noted of higher intensity than the others.

Study 2

Results from Study 2 also support the use of the shortened version of the Ethiopian Lab-TAB (Ethiopian Lab-TAB –S), as most correlation coefficients were significant at the .01 level when related to the data collected in the original Kennedy et al. (2008) study. The exceptions to this finding were one escape variable in Episode 2 and all variables examining either facial or bodily sadness. The low correlations among the sadness variables could be explained by the noted difficulty in recognizing and coding sadness in infants (Rothbart, 2007).

With significant correlations between the original and shortened Ethiopian versions of Lab-TAB, the results from Study 1 could likely be extended here, suggesting that the episodes within the Ethiopian Lab-TAB-S also reliably elicit the temperament behaviors they were designed to. Further implications for the use of this shortened measure will be discussed in greater detail below.

Limitations

A primary limitation to the present study is that data were analyzed post hoc from existing recordings without any new data being collected. This limited the ability to

utilize other temperament measures to assess the validity of the Ethiopian Lab-TAB. Suggestions on overcoming this limitation are discussed below.

A second limitation was the low interrater reliabilities. For the original Ethiopian Lab-TAB data from Shaw et al. (2008), the low interrater reliabilities might be a product of having Ethiopians comprise the original scoring team with Americans performing the reliability checks. For the GTA, interrater reliability ranged from .35 to .72 and for the Ethiopian Lab-TAB-S, from .21 to .78. These values are much lower than the .87 to 1.00 reported by Goldsmith and Rothbart (1991). Lower levels than those previously reported were not entirely unexpected as the present work was a cross-cultural study and thus likely resulted in greater difficulties with coding. This is in part because the Ethiopian infants did not behave in the same manner as American infants would have. In addition, the lower interrater values were mostly found in variables measuring negative affect. Though Hane and colleagues (2006) suggested that negative behaviors may be more easily agreed upon or more consistent across situations than positive behaviors, their work compared the same infants on laboratory assessments and maternal assessments and did not use a cross-cultural sample. And again, Rothbart (2007) notes the greater difficulty in distinguishing between negative affect variables, sadness in particular.

In addition, issues related to the methodology of data collection should be noted as possible explanations for the low interrater reliabilities. The data were collected in an outdoor environment in a portable laboratory that was transported between villages in the Sidama region. This laboratory was constructed of polyvinyl chloride (PVC) pipes with colored bed sheets hung from them. The lens of a video camera was then inserted between two of the sheets to film the infant behaviors. Due to this setting, several

limitations to coding were present. At times, lighting was not conducive to viewing the minute expressions on infants' faces because of lack of sunlight or the direction the camera was focused. At other times, because data were collected within villages, the sounds infants made could not be heard above those made by other villagers or animals within the village (e.g. donkeys, chickens). In addition, the research assistants filming the infants were not skilled videographers and at times lost focus on the infants or did not zoom in to the infants' faces as was needed. All of this contributed to videos that were often of low quality and from which coders at times had to draw inferences regarding what was not visible or heard. Thus, this could be reflected in the interrater reliability values.

As was previously mentioned, studies such as this are of high interest within temperament literature because they are more able to capture the genetic beginnings of temperament before environmental effects have played large roles in further developing the construct (Matheny, Riese, & Wilson, 1985). That being said, the purpose of the original study by Kennedy et al. (2008) was to examine the effects that iron deficiency had on temperament. Thomas, Grant, and Aubuchon-Endsley (in press), in a review of the effects of iron deficiency on cognitive, motor, and affective responding, have noted that the effects of iron deficiency can be both postnatal as well as prenatal. It is unknown, therefore, whether what was measured here and by the original analyses from Kennedy et al. (2008) were a result of nutritional deficiencies experienced prenatally or genetic differences in temperament. As such, conclusions regarding the temperamental attributes of the infants represented here must be considered under this light until additional work can provide greater evidence on this proposed a sensitive period.

Suggestions for Future Research

While replication is of course needed, an additional suggestion for future research would be to collect data that could be analyzed not only by Lab-TAB, but by another temperament assessment to confirm validity. Perhaps an additional parental report measure could be utilized as well. This would aid in overcoming the above noted limitation of having to use the GTA under the constraints of the recordings previously gathered for Lab-TAB. Once validated by other temperament measures with a larger abundance of data on their psychometric properties, the Ethiopian Lab-TAB-S could also be tested for validity in other cultures, making the use of translated materials less necessary, particularly in cultures in which spoken languages have no written text. The use of this laboratory measure might also overcome some of the concerns of parental report assessments of temperament (e.g. Mebert, 1991).

It is further suggested that additional modifications be made to the Ethiopian Lab-TAB-S based on results found here and that further work continue the validation of this measure prior to use. A first suggestion would be to exclude all variables involving sadness, whether bodily or facial, due to the low interrater reliability values. In addition, it might be worth exploring whether higher interrater reliabilities could be achieved by combining fear and anger variables into a "negative affect" category as negative affect in infancy is difficult to separate (Rothbart, 2007). It might also be useful to exclude variables to which a majority of infants fail to demonstrate any behavioral response (e.g. laughter, positive vocalizations). This might serve to increase the median rates of affective variables that do elicit responses from infants, making validity easier to assess

since the affective variables might then have higher scores than the neutral variables of temperament.

Importance

This present study is important because it is adds to the cross-cultural temperament literature and provides a stable assessment base from which temperament can continue to be explored in an Ethiopian sample without necessitating a translation of a parental report measure. As temperament has been shown to be related to multiple important developmental factors such as empathy and conscience (Rothbart, 2007), behavioral problems, personality, neural structures and neurotransmitters, iron deficiency, and general malnutrition, exploring its role cross-culturally is vital.

In addition, the amount of time that would be saved in the large Ethiopian zinc supplementation study scheduled to begin in 2009 is very noteworthy. This longitudinal study plans to examine the effects zinc has on cognitive and temperament development in infants. This project will be a controlled trial with some mothers receiving zinc supplementation during pregnancy, some infants receiving zinc supplements after birth, and some receiving both. This study will involve 900 mother-infant dyads. With that sample size, the difference in using the original Ethiopian Lab-TAB to the shortened version is a difference of 225 to 150 hours – not including time spent coding and in analysis.

Conclusions

The purpose of the present study was to provide evidence of the validity of the Ethiopian Lab-TAB, as well as validity for a shortened version of this cross-cultural measure of temperament. Results from the present study do suggest continued use of

Lab-TAB with an Ethiopian population as well as initial support for the shortened version of this measure, though continued validation is necessary.

It must be stressed that the data presented here are only strong enough to support continued use and exploration of the Ethiopian Lab-TAB-S. The low interrater reliabilities make conclusive reports about the validity of this measure impractical. However, as the first study to assess the use of Lab-TAB in cross-cultural research and with the developmental importance of temperament (Rothbart, 2007), this work makes a large contribution to the cross-cultural temperament literature available.

REFERENCES

Aboud, F. E., & Alemu, T. (1995). Nutrition, maternal responsiveness and mental development of Ethiopian children. *Social Science & Medicine, 41*(5), 725-732.

Bates, J. E. (1989). Concepts and measurements of temperament. In G. A. Kohnstamm, J. E. Bates & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 3-26). New York: John Wiley & Sons.

Bayley, N. (1993). *Bayley Scales of Infant Development: Second edition*. San Antonio, TX: The Psychological Corporation.

Boer, F., & Westenberg, P. M. (1994). The factor structure of the buss and plomin eas temperature survey (parental ratings) in a dutch sample of elementary school children. *Journal of Personality Assessment, 62*(3), 537-551.

Bohlin, G., Hagekull, B., & Lindhagen, K. (1981). Dimensions of infant behavior. *Infant Behavior and Development, 4*, 83-96.

Bornstein, M. H., Gaughran, J. M., & Segui, I. (1991). Multimethod assessment of infant temperament: Mother questionnaire and mother and observer reports evaluated and compared at five months using the Infant Temperament Measure. *International Journal of Behavioral Development, 14*(2), 131-151.

Bridges, L. J., Palmer, S. A., Morales, M., & Hurtado, M. (1993). Agreement between affectively based observational and parent-report measures of temperament at infant age 6 months. *Infant Behavior & Development, 16*(4), 501-506.

Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality*. New York: Wiley.

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Buss, K. A., & Goldsmith, H. H. (1998). Fear and anger regulation in infancy: Effects on the temporal dynamics of affective expression. *Child Development, 69*(2), 359.

Carey, W. B. (1970). A simplified method for measuring infant temperament. *The Journal of Pediatrics, 77*(2), 188-194.

Carey, W. B., & McDevitt, S. C. (1978). Revision of the Infant Temperament Questionnaire. *Pediatrics, 61*(5), 735.

Center, U. o. W. T. (2004). *Lab-TAB*. Retrieved Tuesday, November 27, 2007, from http://psych.wisc.edu/goldsmith/Researchers/GEO/lab_TAB.htm

Cronbach, L. J. (1984). *Essentials of psychological testing* (4th Ed). New York: Harper.

DeVries, M. W., & Sameroff, A. J. (1984). Culture and temperament: Influences on infant temperament in three east African societies. *American Journal of Orthopsychiatry, 544*(1), 83-96.

Gasman, I., Purper-Ouakil, D., Michel, G., Mouren-SimÃ©oni, M. C., Bouvard, M., Perez-Diaz, F., et al. (2002). Cross-cultural assessment of childhood temperament: A confirmatory factor analysis of the french emotionality activity and sociability (eas) questionnaire. *European Child & Adolescent Psychiatry, 11*(3), 101.

Goldsmith, H. H., & Rieser-Danner, L. A. (1990). Assessing early temperament. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational*

assessment of children: Personality, behavior, and context (pp. 245-278). New York: The Guilford Press.

Goldsmith, H. H., & Rieser-Danner, L. A. (1991). Evaluating convergent and discriminant validity of temperament questionnaires for preschoolers. *Developmental Psychology, 27*(4), 566.

Goldsmith, H. H., & Rothbart, M. K. (1988). *The Laboratory Temperament Assessment Battery (Lab-TAB): Locomotor version* (No. 88-01): Oregon Center for the Study of Emotion, technical report.

Goldsmith, H. H., & Rothbart, M. K. (1991). Contemporary instruments for assessing early temperament by questionnaire and in the laboratory. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement* (pp. 249-272). New York: Plenum Press.

Goldsmith, H. H., & Rothbart, M. K. (1999). *The laboratory temperament assessment battery* (Locomotion Version 3.1 ed.). Madison, WI: University of Wisconsin-Madison.

Hagekull, B., Bohlin, G., & Lindhagen, K. (1984). Validity of parental reports. *Infant Behavior and Development, 7*(1), 77-92.

Hane, A. A., Fox, N. A., Cindy, P.-T., Ghera, M. M., & Guner, B. M. (2006). Contextual basis of maternal perceptions of infant temperament. *Developmental Psychology, 42*(6), 1077-1088.

Hubert, N. C., Wachs, T. D., Peters-Martin, P., & Gandour, M. J. (1982). The study of early temperament: Measurement and conceptual issues. *Child Development, 53*(3), 571.

Kennedy, T., Shaw, G., Grant, S., Berhanu, G., Woltamo, T. Roba, A., et al. (2008, April). Growth of breastfed infants in Southern Ethiopia. Poster at the Experimental Biology Conference, San Diego, California.

Kochanska, G., Coy, K. C., Tjebkes, T. L., & Husarek, S. J. (1998). Individual differences in emotionality in infancy. *Child Development, 69*(2), 375.

Lemery, K. S., Goldsmith, H. H., Klinnert, M. D., & Mrazek, D. A. (1999). Developmental models of infant and childhood temperament. *Developmental Psychology, 35*(1), 189-204.

Lerner, J., & Lerner, R. (1983). Temperament and adaptation across life: Theoretical and empirical issues. In P. B. Baltes & O. G. Brim (Eds.), *Life-span development and behavior* (Vol. 5, pp. 197-231). San Diego, CA: Academic Press.

Matheny, A. P., Riese, M. L., & Wilson, R. S. (1985). Rudiments of infant temperament: Newborn to 9 months. *Developmental Psychology, 21*(3), 486-494.

Mathiesen, K. S., & Tambs, K. (2000). The eas temperament questionnaire—factor structure, age trends, reliability, and stability in a norwegian sample. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*(03), 431-439.

Mebert, C. J. (1991). Dimensions of subjectivity in parents' ratings of infant temperament. *Child Development, 62*(2), 352.

Medoff-Cooper, B., Carey, W. B., & McDevitt, S. C. (1993). The early infancy temperament questionnaire. *Journal Of Developmental And Behavioral Pediatrics: JDBP, 14*(4), 230-235.

Pedlow, R., Sanson, A., Prior, M., & Oberklaid, F. (1993). Stability of maternally reported temperament from infancy to 8 years. *Developmental Psychology, 29*(6), 998-1007.

Pettit, G. S., & Bates, J. E. (1984). Continuity of individual differences in the mother-infant relationship for six to thirteen months. *Child Development, 55*(3), 729-739.

Riese, M. L. (1987). Temperament stability between the neonatal period and 24 months. *Developmental Psychology, 23*(2), 216-222.

Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development, 52*(2), 569-578.

Rothbart, M. K. (1986). Longitudinal observation of infant temperament. *Developmental Psychology, 22*(3), 356-365.

Rothbart, M. K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science, 16*(4), 207-212.

Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality & Social Psychology, 78*(1), 122-135.

Sanson, A., Prior, M., Garino, E., & Oberklaid, F. (1987). The structure of infant temperament: Factor analysis of the revised Infant Temperament Questionnaire. *Infant Behavior & Development, 10*(1), 97-104.

Shaw, G., Grant, S., & Thomas, D. (2008, April). Socioeconomic effects of iron deficiency for Ethiopian infants. Poster at the Southwestern Psychological Association Conference, Kansas City, Kansas.

Slabach, E. H., Morrow, J., & Wachs, T. D. (1991). Questionnaire measurement of infant and child temperament. In *Explorations in temperament: International perspectives on theory and measurement* (pp. 205-234).

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/ Mazel.

Thomas, D.G., Grant, S.L., & Aubuchon-Endsley, N. (in press). The role of iron in neurocognitive development. *Developmental Neuropsychology.*

University of Wisconsin Twin Center. (n.d.). *Laboratory Temperament Assessment Battery*. Retrieved January 6, 2008 from http://psych.wisc.edu/goldsmith/researchers/GEO/lab_TAB.htmWorld Health Organization. (2008). *Database on Anaemia.* Retrieved November 11, 2008, from http://www.who.int

APPENDIX A

Tables

Table 1

Summary of the Ethiopian Lab-TAB Protocol

Episode	Name	Temperament Domain	Number of Epochs	Variables Coded
1	Puppet Game	Joy	5	Intensity of Smiling Presence of Laughter Presence of Positive Vocalizations Presence of Positive Motor Acts
2	Parasol Opening	Fear	3	Intensity of Facial Fear Intensity of Vocal Distress Intensity of Bodily Response Intensity of Escape Intensity of Startle
3	Peek-A-Boo	Joy	6	Intensity of Smiling Presence of Laughter Presence of Positive Vocalizations Presence of Positive Motor Acts
4	Task Orientation (Blocks)	Interest	18	Duration of Looking Manipulation of Stimuli Intensity of Facial Interest
5	Toy Retraction	Anger	9	Intensity of Struggle Intensity of Facial Anger Intensity of Distress Vocalizations Intensity of Facial Sadness
6	Basket of Toys	Activity Level	9	Intensity of Manipulation Bouts of Play
7	Person Interest	Interest	6	Duration of Looking Verbal Comments to/about the Person Intensity of Facial Interest
8	Stranger Approach	Fear	7	Intensity of Facial Fear Intensity of Facial Sadness Intensity of Vocal Distress Intensity of Bodily Fear Intensity of Escape
9	Prone/Supine Placement	Activity	6	Intensity of Movement
10	Restraint	Anger	6	Intensity of Facial Anger Intensity of Facial Sadness Intensity of Vocal Distress Intensity of Struggle

Table 2

Interrater Reliability Using Intraclass Correlations

Ethiopian Lab-TAB Episode	Temperament Domain	Episode Description	Median at 6 Months	Median at 9 Months
1	Joy	Puppet Game	.95	.54
2	Fear	Parasol Opening	.42	.58
3	Joy	Peek-a-Boo	.41	.59
4	Interest	Block Episode	.93	.86
5	Anger	Toy Retraction	.72	.63
6	Activity	Basket of Toys	.90	.76
7	Interest	Person Approach	.16	.69
8	Fear	Stranger Approach	.43	.27
9	Activity	Prone & Supine Placement	.80	.73
10	Anger	Restraint	.33	.69

Table 3

Cronbach Alphas from the Ethiopian Lab-TAB

Lab-TAB Episode	Cronbach's Alpha
Episode 1 Intensity of Smiling	.83
Episode 1 Presence of Laughter	.89
Episode 1 Presence of Vocalic Response	.83
Episode 1 Presence of Motor Activity	.83
Episode 3 Intensity of Smiling	.90
Episode 3 Presence of Laughter	.90
Episode 3 Presence of Vocalic Response	.90
Episode 3 Presence of Motor Activity	.90
Episode 4 Duration of Looking	.98
Episode 4 Manipulation of Stimuli	.98
Episode 4 Intensity Facial Interest	.97
Episode 5 Intensity of Struggle	.72
Episode 5 Intensity of Facial Anger	.93
Episode 5 Intensity of Distress Vocalizations	.72
Episode 5 Intensity of Facial Sadness	.93
Episode 6 Intensity of Manipulation	.80
Episode 6 Bouts of Toy Play	.80
Episode 7 Duration of Looking	.94
Episode 7 Verbal Comments To/About Person	.94
Episode 7 Intensity of Facial Interest	.94
Episode 9 Intensity of Movement	.81
Episode 9 Intensity of Movement Prone	.82
Episode 9 Intensity of Movement Supine	.70
Episode 10 Intensity of Facial Anger	.82
Episode 10 Intensity of Facial Sadness	.82
Episode 10 Intensity of Distress Vocalizations	.82
Episode 10 Intensity of Struggle	.82

Table 4

Spearman Brown Prophecy Formula Results from the Ethiopian Lab-TAB

Episode	Number of Original Epochs	Number of Epochs in Shortened Version	Estimated Cronbach's Alpha for Shortened Version	Original Time (in seconds)	New Time (in seconds)
1	5	3	.77	120	72
3	6	3	.82	60	30
4	18	9	.96	180	90
5	9	6	.76	120	80
6	9	6	.72	180	120
9	6	4	.74	60	40
10	6	3	.70	30	15

Table 5

Median Interrater Reliability Values by GTA Variable Using Intraclass Correlations

GTA Episode	GTA Domain	Median Interrater Value
1 - Joy	Joy	.77
Puppet Game	Anger	.20
	Fear	.68
	Interest	.30
	Activity Level	.33
2 – Fear	Joy	.61
Parasol Opening	Anger	*a.*
	Fear	.79
	Interest	.49
	Activity Level	.04
3 – Joy	Joy	.74
Peek-a-Boo	Anger	.19
	Fear	.06
	Interest	.72
	Activity Level	.07
4 – Interest	Joy	.60
Block Episode	Anger	.83
	Fear	1.00
	Interest	.72
	Activity Level	.53
5 – Anger	Joy	.08
Toy Retraction	Anger	.93
	Fear	*a.*
	Interest	.69
	Activity Level	.37
6 – Activity	Joy	.22
Basket of Toys	Anger	.92
	Fear	.28
	Interest	.66
	Activity Level	.51
7 – Interest	Joy	.59
Person Approach	Anger	.70
	Fear	.52
	Interest	.75
	Activity Level	.25
8 – Fear	Joy	.73
Stranger Approach	Anger	.37
	Fear	.76
	Interest	.66
	Activity Level	.04

9 – Activity Prone & Supine Placement	Joy	.50
	Anger	.65
	Fear	.78
	Interest	.38
	Activity Level	.37
10 – Anger Restraint	Joy	.87
	Anger	.81
	Fear	.71
	Interest	.74
	Activity Level	.32

a. Zero variance so no value was calculated.

Table 6

Median GTA Interrater Reliability Values by Episode Using Intraclass Correlations

GTA Episode	Temperament Domain	Episode Description	Median Value
1	Joy	Puppet Game	.33
2	Fear	Parasol Opening	.55
3	Joy	Peek-a-Boo	.19
4	Interest	Block Episode	.75
5	Anger	Toy Retraction	.53
6	Activity	Basket of Toys	.51
7	Interest	Person Approach	.59
8	Fear	Stranger Approach	.66
9	Activity	Prone & Supine Placement	.50
10	Anger	Restraint	.74

Table 7

Descriptive Statistics for GTA Data

Episode	Domain	*N*	Mean	*SD*
Episode 1-Joy Puppet Game	Joy	94	0.32	0.42
	Anger	94	0.01	0.05
	Fear	94	0.04	0.14
	Interest	94	1.77	0.27
	Activity Level	94	0.96	0.28
Episode 2-Fear Parasol Opening	Joy	94	0.05	0.16
	Anger	94	0.05	0.16
	Fear	94	0.82	0.56
	Interest	94	1.67	0.43
	Activity Level	94	0.92	0.26
Episode 3-Joy Peek-a-Boo	Joy	91	0.18	0.38
	Anger	91	0.15	0.48
	Fear	91	0.01	0.09
	Interest	91	1.20	0.55
	Activity Level	91	0.97	0.16
Episode 4-Interest Block Episode	Joy	94	0.00	0.02
	Anger	94	0.01	0.94
	Fear	94	0.00	0.01
	Interest	94	1.61	0.48
	Activity Level	94	1.40	0.39
Episode 5-Anger Toy Retraction	Joy	94	0.01	0.05
	Anger	94	0.22	0.42
	Fear	94	0.00	0.00
	Interest	94	1.24	0.46
	Activity Level	94	1.08	0.24
Episode 6-Activity Basket of Toys	Joy	94	0.01	0.03
	Anger	94	0.08	0.24
	Fear	94	0.01	0.07
	Interest	94	1.76	0.46
	Activity Level	94	1.45	0.41
Episode 7-Interest Person Approach	Joy	93	0.04	0.10
	Anger	93	0.09	0.27
	Fear	93	0.04	0.23
	Interest	93	1.18	0.47
	Activity Level	93	0.99	0.19
Episode 8-Fear Stranger Approach	Joy	94	0.18	0.32
	Anger	94	0.02	0.16
	Fear	94	0.24	0.43
	Interest	94	1.26	0.37
	Activity Level	94	1.02	0.15

Episode 9-Activity Prone & Supine Placement	Joy	93	0.03	0.09
	Anger	93	0.53	0.80
	Fear	93	0.34	0.64
	Interest	93	1.05	0.33
	Activity Level	93	1.09	0.26
Episode 10-Anger Restraint	Joy	89	0.07	0.28
	Anger	89	0.53	0.81
	Fear	89	0.33	0.65
	Interest	89	1.28	0.74
	Activity Level	89	0.78	0.37

a. Possible range from 0 to 3.

Table 8

One-Way ANOVAs for GTA Data

GTA Episode		Sum of Squares	df	Mean Square	F	Sig.
Episode 1	Between Groups	209.27	4	52.32	735.91	< .001
Joy	Within Groups	33.06	465	0.07		
Puppet Game	Total	242.33	469			
Episode 2	Between Groups	173.68	4	43.42	354.42	< .001
Fear	Within Groups	56.97	465	0.12		
Parasol Opening	Total	230.65	469			
Episode 3	Between Groups	106.27	4	26.57	188.00	< .001
Joy	Within Groups	63.59	450	0.14		
Peek-a-Boo	Total	169.86	454			
Episode 4.1	Between Groups	267.11	4	66.78	774.15	< .001
Interest	Within Groups	39.70	460	0.09		
Block Episode	Total	306.79	464			
Episode 4.2	Between Groups	247.02	4	61.75	584.97	< .001
Interest	Within Groups	48.56	460	0.11		
Block Episode	Total	295.58	464			
Episode 4.3	Between Groups	243.75	4	60.94	620.81	< .001
Interest	Within Groups	45.15	460	0.10		
Block Episode	Total	288.91	464			
Episode 5	Between Groups	136.34	4	34.09	379.96	< .001
Anger	Within Groups	41.71	465	0.09		
Toy Retraction	Total	178.05	469			
Episode 6	Between Groups	282.90	4	70.73	807.47	< .001
Activity	Within Groups	40.73	465	0.09		
Box of Toys	Total	323.63	469			
Episode 7	Between Groups	119.00	4	29.75	376.43	< .001
Interest	Within Groups	36.35	460	0.08		
Person Approach	Total	155.35	464			
Episode 8	Between Groups	116.93	4	29.23	313.24	< .001
Fear	Within Groups	43.40	465	0.09		
Stranger Approach	Total	160.33	469			
Episode 9	Between Groups	78.46	4	19.62	79.13	< .001
Activity	Within Groups	114.03	460	0.25		
Prone & Supine	Total	192.50	464			
Episode 10	Between Groups	76.00	4	19.00	51.56	< .001
Anger	Within Groups	162.10	440	0.37		
Restraint	Total	238.06	444			

Table 9

Post hoc Analyses Using Tukeys HSD for GTA Domains

Lab-TAB Episode	Domain 1	Domain 2	Mean Difference	Sig.
Episode 1	Joy	Anger	0.32*	< .001
Joy		Fear	0.28*	< .001
Puppet Game		Interest	-1.44*	< .001
		Activity Level	-0.63*	< .001
	Anger	Fear	-0.03	< .001
		Interest	-1.76*	< .001
		Activity Level	-0.95*	< .001
	Fear	Interest	-1.75*	< .001
		Activity Level	-0.92*	< .001
	Interest	Activity Level	0.81*	< .001
Episode 2	Fear	Joy	0.77*	< .001
Fear		Anger	0.77*	< .001
Parasol Opening		Interest	-0.85*	< .001
		Activity Level	-0.10	0.315
	Joy	Anger	0.00	1.000
		Interest	-1.62*	< .001
		Activity Level	-0.87*	< .001
	Anger	Interest	-1.62*	< .001
		Activity Level	-0.87*	< .001
	Interest	Activity Level	0.75*	< .001
Episode 3	Joy	Anger	0.03	0.985
Joy		Fear	0.17*	0.018
Peek-a-Boo		Interest	-1.02*	< .001
		Activity Level	-0.79*	< .001
	Anger	Fear	0.14	0.079
		Interest	-1.05*	< .001
		Activity Level	-0.81*	< .001
	Fear	Interest	-1.19*	< .001
		Activity Level	-0.10*	< .001
	Interest	Activity Level	0.23*	< .001
Episode 4.1	Interest	Joy	1.63*	< .001
Interest		Anger	1.64*	< .001
Block Episode		Fear	1.64*	< .001
		Activity Level	0.19*	< .001
	Joy	Anger	0.01	1.000
		Fear	0.00	1.000
		Activity Level	-1.44*	< .001
	Anger	Fear	0.00	1.000
		Activity Level	-1.45*	< .001
	Fear	Activity Level	-1.45*	< .001

Episode 4.2	Interest	Joy	1.59*	< .001
Interest		Anger	1.58*	< .001
Block Episode		Fear	1.59*	< .001
		Activity Level	0.21*	< .001
	Joy	Anger	-0.01	1.000
		Fear	0.00	1.000
		Activity Level	-1.38*	< .001
	Anger	Fear	0.01	.999
		Activity Level	-1.37*	< .001
	Fear	Activity Level	-1.38*	< .001
Episode 4.3	Interest	Joy	1.58*	< .001
Interest		Anger	1.55*	< .001
Block Episode		Fear	1.58*	< .001
		Activity Level	0.20*	< .001
	Joy	Anger	-0.02	.983
		Fear	0.01	1.000
		Activity Level	-1.38*	< .001
	Anger	Fear	0.03	.965
		Activity Level	-1.35*	< .001
	Fear	Activity Level	-1.38*	< .001
Episode 5	Anger	Joy	0.21*	< .001
Anger		Fear	0.22*	< .001
Toy Retraction		Interest	-1.02*	< .001
		Activity Level	-0.86*	< .001
	Joy	Fear	0.01	0.999
		Interest	-1.23*	< .001
		Activity Level	-1.07*	< .001
	Fear	Interest	-1.24*	< .001
		Activity Level	-1.08*	< .001
	Interest	Activity Level	0.16*	0.002
Episode 6	Activity Level	Joy	1.44*	< .001
Activity		Anger	1.37*	< .001
Box of Toys		Fear	1.44*	< .001
		Interest	-0.31*	< .001
	Joy	Anger	-0.07	0.470
		Fear	0.00	1.000
		Interest	-1.75*	< .001
	Anger	Fear	0.07	0.506
		Interest	-1.68*	< .001
	Fear	Interest	-1.75*	< .001
Episode 7	Interest	Joy	1.14*	< .001
Interest		Anger	1.09*	< .001
Person Approach		Fear	1.14*	< .001
		Activity Level	0.20*	< .001
	Joy	Anger	-0.06	0.668
		Fear	0.00	1.000
		Activity Level	-0.95*	< .001
	Anger	Fear	0.05	0.716
		Activity Level	-0.89*	< .001
	Fear	Activity Level	-0.94*	< .001

Episode 8	Fear	Joy	0.06	0.714
Fear		Anger	0.21*	< .001
Stranger Approach		Interest	-1.03*	< .001
		Activity Level	-0.78*	< .001
	Joy	Anger	0.16*	0.004
		Interest	-1.08*	< .001
		Activity Level	-0.84*	< .001
	Anger	Interest	-1.24*	< .001
		Activity Level	-1.00*	< .001
	Interest	Activity Level	0.24*	< .001
Episode 9	Activity Level	Joy	1.06*	< .001
Activity Level		Anger	0.56*	< .001
Prone & Supine		Fear	0.75*	< .001
		Interest	0.04	0.983
	Joy	Anger	-0.51*	< .001
		Fear	-0.32*	< .001
		Interest	-1.03*	< .001
	Anger	Fear	0.19	0.066
		Interest	-0.52*	< .001
	Fear	Interest	-0.71*	< .001
Episode 10	Anger	Joy	0.47*	< .001
Anger		Fear	0.20	0.190
Restraint		Interest	-0.75*	< .001
		Activity Level	-0.25*	0.043
	Joy	Fear	-0.26*	0.033
		Interest	-1.21*	< .001
		Activity Level	-0.72*	< .001
	Fear	Interest	-0.95*	< .001
		Activity Level	-0.45*	< .001
	Interest	Activity Level	0.49*	< .001

Table 10

Median Interrater Reliability Values for the Ethiopian Lab-TAB-S Variables Using Intraclass Correlations

Episode	Temperament Domain	Variable	Median Interrater Value
1	Joy	Intensity of Smiling	.77
	Puppet Game	Presence of Laughter	-.05
		Presence of Positive Vocalizations	.40
		Presence of Positive Motor Acts	.36
2	Fear	Intensity of Facial Fear	.05
	Parasol Opening	Intensity of Vocal Distress	.43
		Intensity of Bodily Response	-.01
		Intensity of Escape	a
		Intensity of Startle	.70
3	Joy	Intensity of Smiling	.92
	Peek-a-Boo	Presence of Laughter	.65
		Presence of Positive Vocalizations	.65
		Presence of Positive Motor Acts	-.16
4	Interest	Duration of Looking	.94
	Block Episode	Manipulation of Stimuli	.97
		Intensity of Facial Interest	.56
5	Anger	Intensity of Struggle	.69
	Toy Retraction	Intensity of Facial Anger	.58
		Intensity of Distress Vocalizations	.78
		Intensity of Facial Sadness	.17
6	Activity Level	Intensity of Manipulation	.56
	Box of Toys	Bouts of Play	.84
7	Interest	Duration of Looking	.75
	Person Approach	Verbal Comments to/about the Person	.48
		Intensity of Facial Interest	.62
8	Fear	Intensity of Facial Fear	.12
	Stranger Approach	Intensity of Facial Sadness	.06
		Intensity of Vocal Distress	.91
		Intensity of Bodily Fear	*a.*
		Intensity of Bodily Sadness	-.07
		Intensity of Escape	.48
9	Activity	Intensity of Movement	.76
	Prone & Supine		
10	Anger	Intensity of Facial Anger	.51
	Restraint	Intensity of Facial Sadness	.17
		Intensity of Vocal Distress	.94
		Intensity of Struggle	.62

a. Zero variance so no value was calculated.

Table 11

Median Ethiopian Lab-TAB-S Interrater Reliability Values by Episode Using Intraclass Correlations

Ethiopian Lab-TAB-S Episode	Temperament Domain	Episode Description	Median Value
1	Joy	Puppet Game	.38
2	Fear	Parasol Opening	.24
3	Joy	Peek-a-Boo	.65
4	Interest	Block Episode	.94
5	Anger	Toy Retraction	.67
6	Activity	Basket of Toys	.70
7	Interest	Person Approach	.62
8	Fear	Stranger Approach	.18
9	Activity	Prone & Supine Placement	.76
10	Anger	Restraint	.56

Table 12

Median Pearson rs for the Ethiopian Lab-TAB-S and Ethiopian Lab-TAB Variables

Episode	Temperament Domain	Variable	Correlation Coefficient Value
1	Joy	Intensity of Smiling	.85**
	Puppet Game	Presence of Laughter	.59**
		Presence of Positive Vocalizations	.33**
		Presence of Positive Motor Acts	.60**
2	Fear	Intensity of Facial Fear	.83**
	Parasol Opening	Intensity of Vocal Distress	.99**
		Intensity of Bodily Response	.67**
		Intensity of Escape	-.07
		Intensity of Startle	.62**
3	Joy	Intensity of Smiling	.84**
	Peek-a-Boo	Presence of Laughter	.30**
		Presence of Positive Vocalizations	.74**
		Presence of Positive Motor Acts	.68**
4	Interest	Duration of Looking	.84**
	Block Episode	Manipulation of Stimuli	.90**
		Intensity of Facial Interest	a
5	Anger	Intensity of Struggle	.78**
	Toy Retraction	Intensity of Facial Anger	.66**
		Intensity of Distress Vocalizations	.68**
		Intensity of Facial Sadness	.12
6	Activity Level	Intensity of Manipulation	.86**
	Box of Toys	Bouts of Play	.87**
7	Interest	Duration of Looking	.86**
	Person Approach	Verbal Comments to/about the Person	.58**
		Intensity of Facial Interest	.48**
8	Fear	Intensity of Facial Fear	.70**
	Stranger Approach	Intensity of Facial Sadness	.25
		Intensity of Vocal Distress	.93**
		Intensity of Bodily Fear	.56**
		Intensity of Bodily Sadness	.18
		Intensity of Escape	.33**
9	Activity	Intensity of Movement	.83**
	Prone & Supine		
10	Anger	Intensity of Facial Anger	.79**
	Restraint	Intensity of Facial Sadness	.18
		Intensity of Vocal Distress	.92**
		Intensity of Struggle	.72**

**. Correlation is significant at the 0.01 level (2-tailed).

a. Cannot be computed because at least one of the variables is constant.

Table 13

Median Pearson Rs for the Ethiopian Lab-TAB-S and Ethiopian Lab-TAB Episodes

Lab-TAB Episode	Temperament Domain	Episode Description	Median Value
1	Joy	Puppet Game	.60
2	Fear	Parasol Opening	.67
3	Joy	Peek-a-Boo	.71
4	Interest	Block Episode	.87
5	Anger	Toy Retraction	.67
6	Activity	Basket of Toys	.87
7	Interest	Person Approach	.58
8	Fear	Stranger Approach	.45
9	Activity	Prone & Supine Placement	.83
10	Anger	Restraint	.76

APPENDIX B

Figures

Figure 1

Generalized Temperament Assessment

Episode XXX

Subject number: _____	Scorer: _____
Date scored: _____	

Directions:
Using your own subjective impressions, for each of the 5 epochs for this episode, rate the following temperament aspects on a scale of 0 to 3 by indicating the appropriate number in the chart provided.

Clearly Not Present	A Little Present/ Ambiguous	Somewhat Present	Very Noticeably Present
0	1	2	3

Epochs	1	2	3	4	5	Avg.
Time (Begin/End)						
Joy						
Anger						
Fear						
Interest						
Activity Level						

Figure 2

Ethiopian Version of the Laboratory Assessment Battery

SHORTENED CODING LABTAB DATA
Ethiopia Version, Spring 2008
abridged from full version by Goldsmith & Rothbart

Some Practical General Guidelines For Coding

These comments apply to many of the episodes. They are best understood after becoming familiar with the administration procedures and coding sheets.

1) A single, discrete action that begins during one epoch or trial and extends into the next is coded only in the epoch in which it begins. However, if the action intensifies or changes in any other way, it can be coded again in the next epoch. Note that this guideline is only intended for use in certain instances such as a single, short distress vocalization that has a clear onset and offset. Extended and continuous vocalizations that continue across epoch boundaries should be coded in multiple epochs. Similarly, facial expressions that continue for multiple epochs should be coded in each of the epochs.

2) See Rule 6 below regarding the use of X's.

3) When a trial is curtailed because of experimenter error, camera problems, etc., the missing epochs are marked with an appropriate missing code. (We will used the code "M.") An explanation should be written on the coding sheet. Summary variables are computed as a proportion of the trials that were not missing. If, on the other hand, a trial is curtailed because the infant reaches a point of distress that would prevent further testing, the remaining epochs in the trial are given the same (presumably high) codes achieved in the last valid epoch. Thus, we assume that if the trial had continued, the infant would have continued to cry, struggle, avoid, escape, etc.

4) If a short segment of tape is obscured from view by problems in camera work or unusual movements, it is permissible to make reasonable inferences as to the infant's behavior (e.g. withdrawal) for a few seconds. However, in most cases the epoch should be marked "M" and treated as missing in forming summary variables. An epoch should not be considered missing if the maximum intensity rating was observed during the visible section of the epoch.

5) Coders must guard against becoming either too broad or too literal in interpreting the coding criteria. For instance, on an intensity rating scale, the general descriptors of, for example, "low," "moderate" and "high," should override the illustrative behavioral patterns mentioned in the coding criteria. This must be addressed during training. Consistency both within each subject and between subjects is the most important aspect.

6) Latency is defined as the interval, in seconds, to the first sign of the targeted response. Latency codes of zero are difficult to interpret literally and can also

present problems for mathematical transformation. Therefore, if a response does not occur for which latency is to be measured, the mark of **X** should be entered as a missing data code.

7) Unusual behaviors by the child or mother or mistakes by the camera operator or experimenter should be noted in the margins on the coding sheets.

8) In some coding, such as facial affect measures, the individual characteristics of the child must be taken into account. That is, the fullest smile that one child can show may be less intense than the fullest smile of another child. Although there are obvious dangers in making too great an allowance for such individual differences, they should be a background factor in coding.

9) For most of the episodes, it is not appropriate to begin coding until the child is aware of the stimulus. Epochs where the child is not yet aware of the stimulus (i.e. the stranger) should be considered missing. Similarly, for a response to be coded, it must be directed toward the stimulus. For example, in the Stranger Approach episode, the infant sits on the Experimenter's lap. In this episode, the infant may struggle. If it is clear that this struggle is directed toward the Experimenter and not the Stranger, then struggle should not be coded.

10) Some of the coding sections of the manual contain additional notes or coding conventions. These have been included to help facilitate your judgments of the child's behavior. These are only suggestions and may be changed to suit your needs.

Data Analytic Guidelines

In the future, we expect to provide direct methods for forming composite codes from the raw data contained on the coding sheets. However, the composites that we have derived thus far must be checked in several samples, from our lab and other labs, before they can be recommended for general use. In the meantime, we suggest the following steps in data analysis. We suggest that the initial analysis be at the level of a single episode. The goal of this level of analysis is to derive composites from the episode for use with similar composites from other episodes and data from other assessments.

Steps in analyzing a single Lab-TAB episode:

1) Enter data from coding sheets into an appropriate file format for the statistical program to be used. Data entered should include the raw behavioral counts as well as the latencies, intensities, peak intensities of responses, and averages of various sorts. Some of these parameters might need to be derived from the raw counts and intensities if they are not calculated directly on the coding sheets.

2) Compute descriptive statistics (e.g., mean, SD, minimum, maximum and skewness) for each raw variable. Variables with inadequate variance will be dropped at this point.

3) Plot histograms for all variables. Focusing on the minute by minute averages (latencies, overall durations, peak intensities for the whole episode, average intensities for the episode, etc.), decide if a transformation would help make the distributions more nearly normal. Typically confine transformations to sqrt(x), 1/sqrt(x), and 1/x. If there is difficulty visualizing the effects of these transformations, trial and error will teach quickly. Do not transform unless it helps the distribution substantially; a flat distribution is about as good as a normal one for these analyses. Our experience is that latency scores are frequently best converted to speed scores using the reciprocal of the square root as the appropriate transformation.

4) The next step in analyses of the Lab-TAB measures is the formation of composite variables. First, convert all averages (and any single variable which merits inclusion) into z-scores. Composite variables are formed by intercorrelating all variables (z-scores) within an episode. The variables that correlate significantly and logically fit together can be combined into an appropriately named composite variable. Do not include a significantly correlated variable that seems unjustified theoretically (it's probably a chance correlation). When data from sufficiently large samples are accumulated in our laboratory, factor analytic procedures will be used for this purpose.

5) Note the following information regarding the correlational structure generated for each episode. Often, the correlational structure will be such that a single summary variable can be justified, but there might be clusters of more highly correlated variables within the positive manifold; we refer to these as "component scores." For example, expressive (e.g., facial and vocal) measures and instrumental or motoric measures often fall into different components. It often proves informative to include both the components and the overall summary score in further analyses. Sometimes, component scores will not be intercorrelated, so that no overall summary score is justified.

6) It is important to realize that the data reduction procedure just described does not always capture the temporal variability in infant responses very well. That is, latency scores are combined with other parameters so that their independent effect is not reflected in the component or summary scores. Thus, investigators might choose to carry latency/speed scores into later analyses. We are currently engaged in extensive study of the temporal patterning of infant responses. This obviously touches on the issue of emotion regulation, and extensive work on this topic, within the Lab-TAB episodes, is underway in the Wisconsin laboratory.

7) Both the lower order and higher order composites can be moved to other data files for combination with data from temperament questionnaires,

interviews and other episodes.

***Note:** A danger exists in dealing with missing data. When a variable that enters into a composite is missing, different statistics programs have different default options for dealing with the situation. Our rule of thumb is to compute a higher-order composite if more than half of the lower order constituents are non-missing. This treatment of missing data can be accomplished in several ways that all involve a logical prediction of what the subject's response would have been (Little & Rubin, 1987). It is particularly important that subjects whose testing in an episode is terminated due to distress or other extreme reactions not be deleted from the analysis.

Coding from CDs

- Most of the CDs have 3 video files on them. For example, for subject number 13, these will be labeled:
  - **Fe013Restraint.mpg**, which has the restraint episode taken during the length measurement. It is episode number 10 in the listing below;
  - **Fe013A.mpg**, which has episodes 1-5;
  - **Fe013B.mpg**, which has episodes 6-9.
- Some of the CDs we burned later have files from several babies on them. For each baby, there will be restraint episode file (e.g., **Fe083Restraint.mpg**) and only one file with all other LabTab episodes (e.g., **Fe083A&B**)

- The beginning and end of each episode are marked by a beep of the stopwatch and the camera operator saying "start" or "end." (They might also say "begin" and/or "stop.")
  - The beep of the stopwatch should be used as the time marker. Only use the word (start, end, begin, or stop) if no beep is heard.
  - Important time marks within an episode (e.g., marking the 15-second intervals in the Toy Retraction episode) are typically marked with double beeps.
  - If no beep or word is heard at the beginning of an episode, start coding as soon as the episode begins on the CD.
  - If no beep or word is heard at the end of an episode, stop coding when the proper amount of time has passed. These times are described under the **Scoring** paragraph for each episode.
  - If the camera operator continues to time an episode longer than it should be, stop coding when the proper amount of time has passed.

- We have found it best to watch an entire episode all the way through before coding. That way a general understanding of the subject can be gained as well as noticing approximately when particular behaviors occur.

- If you hear a baby crying or making other sounds in an episode, do not assume that it is the baby being tested. There were many babies in the area, so if you are coding crying, be sure that it is the baby you are looking at and not one off-camera.

1. EC 3,2. PUPPET GAME (Joy/Pleasure)

Rationale

This episode measures enjoyment in response to social stimulation. The use of puppets allows a more standardized social interaction than is possible when the mother has a prominent role.

Scoring

The episode begins with the appearance of the puppets from under the table and is divided into 3 trials for scoring. The first trial begins when the puppets appear from under the table, until the puppets make contact with C at the beginning of the first tickle. The second trial begins when the puppets make contact with C for the first tickle until the puppets touch C for the second tickle. The third trial begins when the E puts the puppets on the table in front of C and then continues for 30 s. The epochs are coded by indicating the occurrence of the specified behavior, or by rating the intensity of the behavior. When an intensity rating is requested, the highest intensity observed should be coded.

Variables to be coded: a) Latency to joy. b) Intensity of smiling. c) Laughter. d) Positive vocalizations. e) Positive motor acts. f) Latency to approach. g) Engagement with toy. h) Parent behavior. i) Baseline state.

Definitions of variables:

a) Latency to joy: Interval, in seconds, from the start of the episode to the first sign of joy (facial, postural, or vocalic). Facial joy can be the joy marker here if b) Intensity of smiling is coded > 0; postural joy can be the joy marker here if e) positive motor activity is coded > 0; vocalic joy can be the joy marker here if d) positive vocalization is coded > 0.

b) Intensity of smiling: Peak intensity of facial joy is noted in each epoch using AFFEX (See Appendix A for definitions) and rated on the following scale: 0= No smiling at all. 1= Small smile, with lips slightly upturned, and no involvement of cheeks or eyes. 2= Medium smile, with lips upturned, perhaps mouth open, slight bulging of cheeks, and perhaps some crinkling about the eyes. 3= Large smile, with lips stretched broadly and upturned, perhaps mouth open, definite bulging of cheeks and noticeable crinkling of eyes.

c) Laughter: Presence of laughter in each of the epochs is noted; laughter should be more intense than positive vocalizations and usually has a rhythmic quality. 0 = Not present. 1 = Present.

d) Positive vocalizations: Presence of positively toned babbling, squealing, and similar behaviors in each epoch is noted. 0= Not present. 1= Present.

e) Positive motor activity: Positive motor acts include the following: banging of hands on table; clapping of hands; waving arms in excitement; attempts, reaches

for and/or plays with puppets during dialogue; definite leaning forward, or rocking, towards the puppets. 0= Not present. 1= Present.

f) Latency to approach: Interval, in seconds, from E's laying the puppets on the table in front of C to C's first contact with them.

g) Engagement with toy: the child's level of engagement with the toy should be noted. 0= Indifferent to the toy. 1= Neutral reaction to toy, looks at toy with mild interest. 2= Fully engaged with toy; likes toy, engrossed in toy.

***Note:** When considering engagement with the toy, remember that staring, leaning and reaching are equally important, especially for infants who may not have the motor skills required to reach for a moving toy. Consider the intensity of the stare: Is it a blank stare or is the child mentally engaged? A child should not automatically lose engagement points just because s/he does not reach.

h) Parent behavior 0= Parent verbally or physically (i.e. bouncing) attempts to elicit joy from C. 1= Mild interference; 1-2 comments directed at C or adjustments of C. Snapping the fingers or other similar attempts to get the infant's attention would be coded as 1. These
comments or adjustments are not intended to elicit joy. 2= Not interfering, neutral.

i) Baseline state: The child's state prior to the beginning of an episode:
1= drowsy. 2= alert/calm. 3= alert/active. 4= fussy (upset but not crying). 5= crying.

2. EC 1,1. PARASOL OPENING (Fear)

Rationale

A parasol opens fairly rapidly and without warning as the infant sits in an enclosed booth. Substantial research indicates that looming visual stimuli evoke startle reactions; however, the relation of startle and more conventional fear measures is less clear. Some infants also exhibit signs of fearfulness in this episode. Elements of both novelty and intrusiveness are present. Repeated trials allow evaluation of changes in reactivity.

Coding

This episode consists of three, 5 s trials, each trial beginning when the parasol begins to open. E rapidly opens the parasol, then slowly closes the parasol, taking 5 s. This process is repeated twice. Uncoded intervals of 5 s separate the three trials. In this episode, if the child is not looking in the general direction of the parasol when it is opened, it is marked as M, regardless of the responses. This rule is applied to each trial separately, not just to the first one.

Variables to be coded: a) Latency to first fear response. b) Intensity of facial fear.c) Intensity of distress vocalizations. d) Intensity of bodily fear. e) Intensity of escape behaviors. g) Intensity of startle. h) Baseline state. i) Parent behavior.

Definitions of variables:

a) Latency to first fear response: Interval, in seconds, from the start of the trial to the first definite fear response (can be facial or vocalic). A definite fear response is any facial or vocal response that would be coded as a '1' or higher on the coding sheet. If b) Intensity of facial fear is scored as 1 or above, this indicates a facial fear response. If c) Intensity of distress vocalizations is scored as 1 or above, this indicates vocalic fear.

b) Intensity of facial fear: Presence of fear or fear blends using AFFEX (see appendix A
for more information): 0 = No facial region shows codable fear movement. 1 = Only one facial region shows codable movement, identifying a low intensity fear, or expression is ambiguous. 2 = Only 2 facial regions show codable movement, or expression in one region is definite. 3 = An appearance change occurs in all 3 facial regions, or coder otherwise has impression of strong facial fear.

***Note**: Non-fear expressions may occur in this episode. Specifically, it is common to see lip corners drawn straight back along with the inner corners of the brows drawn down and together, eyes squinted, and cheeks raised. Lip corners drawn straight back is usually associated with fear while the rest of this expression (brows, cheeks, and eyes) is usually associated with anger. To guard against including these movements in our fear coding, we use the following convention. If a fear mouth (corners drawn straight back) occurs with anger brows, cheeks, and eyes, the highest possible facial

fear intensity rating is '1.' Another possibility with the above facial configuration is that the mouth is a low intensity anger mouth and therefore the entire configuration is anger. If the lip corners are drawn straight back and are beginning to look squarish or the lips are pressed tightly together while being drawn back you may want to consider a '0' for fear.

c) Intensity of distress vocalizations*: Peak intensity of distress vocalizations is noted in
each epoch and rated on the following scale: 0 = No distress. 1 = Mild vocalization that may be difficult to identify as hedonically negative. 2 = Definite whimpering, limited to a short (1-2 second) duration. 3 = Longer whining, fussing, mild protest, or low-intensity cry (cry has
extended or rhythmic quality). 4 = Definite non-muted crying. 5 = Full intensity cry/scream (child is losing control).

***Note:** some vocalizations in this episode will not be fear-related and should not be coded

d) Intensity of bodily response: Peak intensity of bodily response is noted in each trial and
rated on the following scale: 0 = No sign of bodily fear. 1 = Decreased activity: an apparent and sudden decrease in the activity level of C. 2 = Tensing: visible and sustained tensing of the muscles, associated with decreased activity. 3 = Trembling due to extreme muscular tension.

e) Intensity of escape*: Peak intensity of escape is noted in each epoch and rated on the
following scale: 0 = No escape behavior or social referencing. 1 = Mild or fleeting escape behavior (e.g. turning away including social referencing the mother, sinking into chair). 2 = Moderate escape behavior resulting in significant, but not extreme attempts to get away or resist. Full body movements such as arching back, twisting away, and leaning away are included as well as hitting, pushing, and/or slapping. 3 = Vigorous escape behavior, often lasting for most of the epoch (e.g. twisting away, leaning away, hitting, pushing, and/or slapping). Here social referencing may or may not occur.

***Note**: Escape is a very active behavior; it should be coded only when the action of escaping is made. It should only be coded when the initial escape behavior is made or when it is repeated or intensified (see #1 of "some practical general guidelines for coding" for further information).

f) Intensity of startle: The peak intensity of startle is noted in each epoch and rated on the following scale: 0 = No startle. 1 = Very mild or ambiguous startle reaction; perhaps little more than an eye blink. 2 = Moderate startle reaction, e.g., blink and head movement. 3 = Full and extended startle reaction, e.g., blink and body movement.

Intensity of bodily response can be coded as a part of the startle response. However, Intensity of facial fear, Intensity of distress vocalizations (and vocal fear) and Intensity of escape must follow the termination of the startle response.

g) Baseline state: 1= tired/drowsy. 2= alert/calm. 3= alert/active. 4= fussy. 5= crying.

h) Parent behavior 0= Interfering; emotionally loaded statements to C, soothing, reprimanding C, commanding, or generally disrupting. 1= Mild interference; 1-2 comments directed at C or adjustments of C. These comments or adjustments are not emotionally loaded. 2= Not interfering, neutral.

3. EC 3,4. MODIFIED PEEK-A-BOO GAME (Joy/Pleasure)

Rationale

This episode measures pleasure in response to social stimulation. The episode has ample precedent in developmental research. The operationalization of the peek-a-boo game in Lab-TAB minimizes the variance that might be attributable to maternal behavioral differences, since mothers follow a script in this version of peek-a-boo.

Scoring

The episode begins with E's first "Where's Mommy?" and consists of 3 trials, each one beginning with "Where's Mommy?" The first trial will end with the beginning of the next "Where's Mommy?" The last trial ends 5 s after it begins. The trials are coded by indicating the occurrence of the specified behavior, or by rating the intensity of the behavior. When an intensity rating is requested, the highest intensity observed should be coded. Noticing the mother needs to only occur one time, not on each trial.

Variables to be coded:

a) Latency to joy. b) Intensity of smiling. c) Laughter. d) Positive vocalization. e) Positive motor activity. f) Effectiveness of parent. g) Baseline state.

Definitions of variables:

a) Latency to joy: Interval, in seconds, from the start of the trial to the first sign of joy (facial, postural, or vocalic). Facial joy can be the joy marker here if b) Intensity of smiling is coded > 0; postural joy can be the joy marker here if e) positive motor activity is coded > 0; vocalic joy can be the joy marker here if d) positive vocalization is coded > 0.

b) Intensity of smiling: Peak intensity of facial joy is noted in each epoch using AFFEX (See Appendix A for definitions) and rated on the following scale: 0= No smiling at all. 1= Small smile, with lips slightly upturned, and no involvement of cheeks or eyes. 2= Medium smile, with lips upturned, perhaps mouth open, slight bulging of cheeks, and perhaps some crinkling about the eyes. 3= Large smile, with lips stretched broadly and upturned, perhaps mouth open, definite bulging of cheeks and noticeable crinkling of eyes.

c) Laughter: Presence of laughter in each trial is noted; laughter should be more intense than positive vocalizations and usually has a rhythmic quality. 0= Not present. 1= Present.

d) Positive vocalization: Presence of positively toned babbling, squealing, etc. is noted in each trial . 0= Not present. 1= Present.

e) Positive motor activity: Presence of banging of hands on table, clapping, waving of arms in excitement, reaching toward the doors, leaning toward or rocking forward, etc. is noted in each trial. 0= Not present. 1= Present.

f) Effectiveness of parent: the parent's effectiveness as a participant in the episode. The parent gets one overall code for the entire episode. 0= Ineffective; does not say Peek-A-Boo or says Peek-A-Boo more than two times, overly enthusiastic or no enthusiasm at all. 1= Mildly effective; tone is either somewhat too positive or too flat/negative. 2= Effective, says Peek-A-Boo once in a positive tone.

g) Baseline state: The child's state prior to the beginning of the episode: 1= drowsy. 2= alert/calm. 3= alert/active. 4= fussy. 5= crying.

4. EC 4,1. TASK ORIENTATION (BLOCKS) (Interest/Persistence)

Rationale

This episode provides an opportunity for the child to manipulate a set of blocks. Blocks can facilitate a wide variety of responses. All children are capable of many of these responses; therefore, the primary determinant of differences in amount of manipulation of the blocks is motivation. In this episode, motivation is equated with the emotion of interest and, in particular, with its duration parameter, persistence.

Scoring

This episode lasts 90 seconds and is subdivided into nine 10 second epochs. The epochs are coded by indicating the occurrence of the specified behavior, or by rating the intensity of the behavior. When an intensity rating is requested, the highest intensity observed should be coded. Child may be determined to notice the blocks before the start signal is given.

Variables to be coded: a) Intensity of facial interest. b) Duration of looking. c) Latency to look away. d) Manipulation of stimuli. e) Parent behavior. f) Baseline state.

Definitions of variables:

a) Intensity of facial interest*: Peak intensity of facial interest is noted in each epoch using AFFEX (See Appendix A for definitions) and rated on the following scale: 0= No facial region shows codable interest, infant is not looking at the blocks. 1= Codable interest, identifying a low intensity interest, C may be simply attending to the blocks. 2= A definite facial indication of interest occurs or coder otherwise has impression of strong facial interest. C's mouth may fall open; C's eyebrows raise straight up; eyes widen; or eyebrows are down and together like in concentration.

b) Duration of looking: The amount of time C spends looking at the blocks. This is recorded for each epoch. Ratings are made on the following 4-point scale: 0= Does not look at blocks at all. 1= 1-4 seconds. 2= 5-8 seconds. 3= 9-10 seconds.

c) Latency to look away: Interval, in seconds, from the start of the episode to the first look away from the stimuli.

d) Manipulation of stimuli: Time spent touching, holding, and mouthing the blocks within each epoch. This does not include throwing the blocks off the table. Ratings are made on the following 4-point scale: 0= Does not manipulate blocks at all. 1= 1-4 seconds. 2= 5-8 seconds. 3= 9-10 seconds.

e) Parent behavior 0= Interfering; parent actively encourages or discourages child to attend to or manipulate the blocks. 1= Mild interference; 1-2 comments directed at C or adjustments of C. These comments or adjustments do not directly

encourage/discourage the child's involvement with the blocks. For codes of either 0 or 1, the mother may also interfere by handing the blocks to the child. 2= Not interfering, neutral.

f)Baseline state: The child's state prior to the beginning of the episode:1= drowsy. 2= alert/calm. 3= alert/active. 4= fussy. 5= crying.

5. EC 2,3. TOY RETRACTION (Anger/Frustration)

Rationale

This episode provides an opportunity for the expression of anger by interrupting the exploration of a toy. The context is interpersonal, and the situation can be viewed as a violation of social norms.

Coding

This episode consists of two, 15-second trials--each presentation of the toy is one trial. For coding purposes, each trial begins when M starts to take the toy from C – that is, when the mother starts to reach for the toy – and ends 15 seconds later. Each trial is divided into three, 5-second epochs.

Variables to be coded: a) Latency to first anger response. b) Latency to first sadness response. c) Intensity of struggle. d) Intensity of facial anger. e) Intensity of distress vocalizations.
f) Interest in toy. g) Parent effectiveness. f) Baseline state.

Definition of variables:

a) Latency to first anger response: Time, in seconds, starting when M removes toy from C to the first sign of anger (facial, vocalic, postural, or instrumental). Facial anger can be the anger marker here if d) Intensity of facial anger is coded > 0; vocalic anger can be the anger marker here if e) intensity of distress vocalizations is coded as 3, 4, or 5 (meaning a cry).

b) Latency to sadness response: interval, in seconds, to the first sign of sadness (facial, postural, or vocalic).

c) Intensity of struggle: Peak intensity of struggling in each epoch is rated on the following scale: 0 = No struggling to hold on to the toy at all. No resistant movement. 1 = Low intensity struggle. 2 = Medium intensity struggle. Sporadically pulls toy away from parent using
arms and/or body. Movements could include medium intensity pulling of toy, leaning forward, arching back or kicking. Generally lasts 2 - 3 s. 3 = Moderately high intensity struggle. Near continuous moderate intensity pulling of toy. Can include the same movements as number 2 with higher intensity. Generally lasts 3 - 4 s. 4 = High intensity struggle. Continuous movement of moderately high intensity with intervals of high intensity movements to get toy. Generally lasts 5 or more s. *Struggling includes behavior directed toward the toy after it has been removed from the child's hand (e.g, reaching or attempting to crawl after the toy)*

d) Intensity of facial anger: Presence of anger or anger blends is noted in each epoch using AFFEX (See Appendix A for definitions) and rated on the following scale: 0 = No facial region shows codable anger movement. 1 = Only one facial region shows codable movement, identifying a low intensity anger, or expression is ambiguous. 2 = Only 2 facial regions show codable movement, or expression in one region (e.g.,

brows) is definite. 3 = An appearance change occurs in all 3 facial regions, or coder otherwise has impression of strong anger.

e) Intensity of distress vocalizations*: Peak intensity of distress vocalizations is noted in
each epoch and rated on the following scale: 0 = No distress. 1 = Mild protest verbalization that may be difficult to identify as hedonically negative. 2 = Definite protest, limited to a short (1-2 second) duration. 3 = Longer protest, fussing or mild, low-intensity cry (cry has extended or
rhythmic quality). 4 = Definite non-muted crying. 5 = Full intensity cry/scream (child is losing control).

***Note**: some vocalizations in this episode will not be anger-related and should not be coded.

f) Interest in toy: 0 = Merely holding toy without attending to it or shaking it.. 1 = One behavior showing moderate interest, e.g., just shaking, looking or mouthing. 2 = Intense interest such as showing two behaviors, e.g., shaking and mouthing.

e) Effectiveness of parent: the parent's effectiveness as a participant in the episode is coded. The parent receives one overall code for the entire episode. 0= Not effective: Parent does not move toy out of C's reach or returns it prematurely in two out of the three trials. 1= Mildly effective: Parent does not move toy out of C's reach or returns it prematurely in one of the trials. 2= Effective: Parent follows the instructions that E has given in all three of the trials.
f) Baseline state: The child's state prior to the beginning of an episode:1= tired/drowsy. 2= alert/calm. 3= alert/active. 4= fussy. 5= crying.

6. EC 5,3. BASKET OF TOYS (Activity Level)

Rationale

This episode provides a measure of activity level during object-oriented play. The child's trunk is supported so that maintaining access to the toys is not a problem. Observed individual differences should reflect the rate and pattern of reaching and manipulation.

Coding

The episode lasts 2 minutes. The 2 minutes are divided into 20 second epochs for coding purposes.

Variables to be coded: a) Latency to approach any toy. b) Intensity of toy manipulation. c) Bouts of toy play. d) Parent behavior. e) Baseline state.

Definitions of variables:
a) Latency to approach any toy: Interval, in seconds, from the beginning of the episode until the child touches the first toy (minimum of 1 sec).

b) Intensity of manipulation: How intensely the child manipulates toys is rated on the following 5 point scale: 0= Little or no manipulation: looking at toys without touching any. 1= Low level of manipulation: playing with, touching, or mouthing toys. (this code does not include shaking toys). 2= Moderate level of manipulation: shaking toys, repetitively squeezing toys, low intensity banging or C plays with 5 or more toys at low intensity, C plays with one toy for the entire epoch at low intensity. 3= High manipulation: Higher intensity '2's that are near continuous and last for approximately half the epoch. 4= Extremely high manipulation: includes movements from '2' but at a higher intensity; banging and/or throwing of toys in a rough manner. Movements are continuous in nature and last for almost the entire epoch.

Note: If there are no new bouts of toy play, but the child is still manipulating the toy or toys from the previous epoch, manipulation energy should still be scored.

c) Bouts of toy play: Number of different bouts of toy play during each epoch. A new bout is usually signaled by picking up or playing with a new toy. In some cases, a new bout may involve a totally different type of activity with the same toy. Sucking is counted only once per epoch and from then on sucking is counted as part of whatever manipulation C is doing (e.g. sucking and shaking is considered to be one bout of toy play). Also, all touching has to be intentional; if holding a toy during the transition between epochs, just count bout in the first epoch it appeared.
Some helpful hints for coding bouts of toy play: The basket is counted as a toy.
Do not count rapid, multiple touches: e.g. C is hitting or pulling on the same toy repeatedly.
If C is still playing with a toy when a new epoch starts you do not code it as a new bout of toy play.

d) Parent behavior 0= Interfering; emotionally loaded statements to C, soothing, reprimanding C, commanding, or generally disrupting. 1= Mild interference; 1-2 comments directed at C or adjustments of C. These comments or adjustments are not emotionally loaded. 2= Not interfering, neutral.

e) Baseline state: The child's state prior to the beginning of the episode: 1= tired/drowsy. 2= alert/calm. 3= alert/active. 4= fussy. 5= crying.

7. EC 4,2. PERSON INTEREST (Interest/Persistence)

Rationale

In contrast to episodes using inanimate stimuli, this episode examines interest in a social context where a female experimenter acts out a scripted set of behaviors in the presence of the child. The child's interest in the person's activities is measured.

Scoring

The episode continues for 60 - 70 s and is divided into 10 second epochs for scoring purposes. The first epoch begins with the first word of the conversation. The coded portion of the conversation is 30 s (3 epochs). Double beeps will mark the end of 30 seconds and the person will cease talking and stands; this also marks when the fourth epoch begins. The fifth epoch is 10 s long. The sixth epoch begins as soon as the fifth concludes and is also 10 s long. Any behavior beyond this time frame is uncoded. Because there are no auditory codes between the end of the talking and the end of the episode, the time of the end of the episode must be found and 20 seconds subtracted from that time. This is when the fourth epoch ends and the fifth begins. The epochs are coded by indicating the occurrence of the specified behavior, or by rating the intensity of the behavior. When an intensity rating is requested, the highest intensity observed should be coded.

Variables to be coded: a) Intensity of facial interest. b) Duration of looking. c) Latency to look away. d) Vocalizations about E. e) Parent behavior. f) Baseline state.

Definitions of variables:

a) Intensity of facial interest*: Peak intensity of facial interest is noted in each epoch using AFFEX (See Appendix A for definitions) and rated on the following scale: 0) No facial region shows codable interest; infant is not looking at the unfamiliar experimenter. 1) Codable interest, identifying a low intensity interest; C may be simply attending to the unfamiliar person.
2) A definite facial indication of interest occurs or coder otherwise has impression of strong facial interest. C's mouth may fall open; C's eyebrows raise straight up; eyes widen; or eyebrows are down and together like in concentration.

b) Duration of looking: The amount of time C spends looking at the unfamiliar experimenter. Ratings are made on the following 4-point scale: 0) Does not look at unfamiliar experimenter at all. 1) 1-4 seconds. 2) 5-8 seconds. 3) 9-10 seconds.

c) Latency to look away: Interval, in seconds, from the start of the episode to the first look away from the unfamiliar experimenter.

d) Vocalizations: Vocalizations that indicate an interest in the unfamiliar experimenter. The presence or absence of vocalizations are noted within each epoch. 0= absent. 1= present.

e) Parent behavior 0= Interfering; parent actively encourages or discourages child to attend to the unfamiliar person. 1= Mild interference; 1-2 comments directed at C or adjustments of C. These comments or adjustments do not directly encourage/discourage the child to attend to the unfamiliar person. 2= Not interfering, neutral.

f)Baseline state: The child's state prior to the beginning of an episode:1= drowsy. 2= alert/calm. 3= alert/active. 4= fussy. 5= crying.

8. EC 1,3. STRANGER APPROACH (Fear)

Rationale

An adult male stranger will approach and stare at the child in a standardized fashion. The elements of novelty and intrusiveness should elicit various degrees of fearful distress and
avoidance.

Scoring

This episode consists of 3 stages. Stage one is divided into three epochs. The first epoch of stage one begins with the entry of the male stranger and includes the initial pause (about 10 s). The second epoch is S's approach to the half-way point (about 10 s). The third epoch is the pause in which S speaks to C (about 5 s). The first two epochs are calculated by measuring the time from the initial auditory start signal to when the stranger begins to speak. The length of each of the first two epochs are thus this interval divided by 2. For example, if the duration of the interval between the start signal and the stranger beginning to speak is 18 seconds, then the first two epochs are each 9 seconds long. The second and third stages are divided into two epochs each. The fourth epoch is S's approach to within .3 m of C (about 10 s). The fifth epoch is the following pause (about 10 s). The fourth and fifth epochs are calculated by measuring the time from the time the stranger stops speaking to when the stranger begins to pick up the child. The length of each of the fourth and fifth epochs are thus this interval divided by 2. The sixth epoch is when S picks up and holds C (about 15 s), and the seventh epoch is when S replaces C in the high chair (about 10 s). If there are less than 4 seconds in this last epoch, do not score but code as unmarked (M). The peak intensity of the following variables is scored within each epoch, as well as the presence/absence of escape behaviors and bodily sadness.

Variables to be coded: a) Latency to first fear response. b) Latency to first sadness response.
c) Intensity of facial fear. d) Intensity of facial sadness. e) Intensity of distress vocalizations. f) Intensity of bodily fear. g) Presence of bodily sadness. h) Intensity of escape. i) Baseline state. j) Parent behavior.

Definitions of variables:
a) Latency to first fear response: Interval, in seconds, from the start of the episode to the first definite fear response (facial, vocalic, and postural which includes bodily fear or escape behavior). A definite fear response is any response that would be coded as a '1' or higher on the coding sheet. If c) Intensity of facial fear is scored as 1 or above, this indicates a facial fear response. If e) Intensity of distress vocalizations is scored as 1 or above, this indicates vocalic fear. If f) Intensity of bodily fear is scored as 1 or above, this indicates postural fear.

b) Latency to sadness response: First sign of sadness (facial, postural, or vocalic).

c) Intensity of facial fear: Peak intensity of facial fear or fear blends is noted in each epoch using AFFEX (See Appendix A for definitions) and rated on the following

scale: 0= No facial region shows codable fear movement. 1= *Only one facial region shows codable movement, identifying a low intensity fear, or expression is ambiguous. 2= Only 2 facial regions show codable movement, or expression in one region is very clear. 3= An appearance change occurs in all 3 facial regions, or coder otherwise has impression of strong facial fear.

***Note**: Non-fear expressions may occur in this episode. Specifically, it is common to see lip corners drawn straight back along with the inner corners of the brows drawn down and together, eyes squinted, and cheeks raised. Lip corners drawn straight back is usually associated with fear while the rest of this expression (brows, cheeks, and eyes) is usually associated with anger. To guard against including these movements in our fear coding, we use the following convention. If a fear mouth (corners drawn straight back) occurs with anger brows, cheeks, and eyes, the highest possible facial fear intensity rating is '1.' Another possibility with the above facial configuration is that the mouth is a low intensity anger mouth and therefore the entire configuration is anger. If the lip corners are drawn straight back and are beginning to look squarish or the lips are pressed tightly together while being drawn back you may want to consider a '0' for fear.

d) Intensity of facial sadness: Peak intensity of facial sadness or sadness blends is noted in each epoch using AFFEX (See Appendix A for definitions) and rated on the following scale: 0= No facial region shows codable sadness movement. 1= Only one facial region shows codable movement; identifying very low or ambiguous facial sadness. 2= Only 2 facial regions show codable movement, or expression in one region is very clear. 3= An appearance change occurs in all 3 facial regions or coder otherwise has impression of strong facial sadness.

e) Intensity of distress vocalizations*: 0= No distress. 1= Mild vocalization that may be difficult to identify as hedonically negative. 2= Definite whimpering, limited to a short (1-2 second) duration. 3= Longer whining, fussing, mild protest, or low-intensity cry (cry has extended or
rhythmic quality). 4= Definite non-muted crying. 5= Full intensity cry/scream (child is losing control).

***Note**: some vocalizations in this episode will not be fear-related and should not be coded.

f) Intensity of bodily fear: Peak intensity of bodily fear is noted in each epoch and rated on the following scale*: 0= No sign of bodily fear. 1= Decreased activity: an apparent and sudden decrease in the activity level of C. 2= Tensing: visible and sustained tensing of the muscles, associated with decreased activity. 3= Trembling due to extreme muscular tension.

***Note**: Bodily fear should only be coded across epochs when the intensity is ‘2’ or higher, unless a lower intensity behavior, '1', is repeated or changes in the following epochs (see general coding guidelines #1).

g) Presence of bodily sadness: (slight slump, drop of head, slumped shoulders, head in arms or hands) 0= No detectable sadness. 1= Very clear, detectable sadness.

h) Intensity of Escape*: Peak intensity of escape is noted in each epoch and rated on the following scale: 0= No escape behavior or social referencing. 1= Mild or fleeting escape behavior (e.g. turning away including social referencing the mother, sinking into chair). 2= Moderate escape behavior resulting in significant, but not extreme attempts to get away or resist. Full body movements such as arching back, twisting away, and leaning away are included, as well as hitting, pushing, and/or slapping. 3= Vigorous escape behavior, usually involving linked, intense full-body movements like those found in '2'. These movements usually last for the entire epoch.

***Note**: Escape is a very active behavior; it should be coded only when the action of escaping is made. Also, it should only be coded when the initial escape behavior is made or when it is repeated or intensified (See #1 of "Some Practical General Guidelines for Coding" for further information).

i) Baseline state: 1= tired/drowsy. 2= alert/calm. 3= alert/active. 4= fussy. 5= crying. *In some cases the mother must hold the infant because he/she becomes very upset when sitting on the E's lap. Here it is assumed that the baseline state is 4 (fussy), unless the baby continues to cry (but the procedure should have been halted in this case).*

j) Parent behavior: 0= Interfering; emotionally loaded statements to C, soothing, reprimanding C, commanding, or generally disrupting. 1= Mild interference; 1-2 comments directed at C or adjustments of C. These comments or adjustments are not emotionally loaded. 2= Not interfering, neutral.

9. EC 5,2. PRONE & SUPINE PLACEMENT (Activity Level)

Rationale

A measure of a child's activity during a situation where no overt activity-eliciting stimuli are present is provided by alternatively placing a child in the supine and prone positions.

Coding

The episode is divided into two, 20-second trials. Each trial is further divided into two, 10-second epochs.

Variables to be coded: a) Latency to move once placed on the mat. b) Latency to turn over. c) Intensity of movement. d) Parent behavior. e) Baseline state.

Definitions of variables:

a) Latency to move once placed on the mat: Interval, in seconds, from the beginning of the episode to the first occurrence of movement on the mat. Latency to move must involve movement that can be coded as 1 or greater under c) Intensity of movement.

b) Latency to turn over once placed on the mat: Interval, in seconds, from the beginning of the episode to C turning over. The infant must turn all the way over; latency should be measured up to when the turn is complete.

c) Intensity of movement*: The intensity of movement is rated for each interval on a scale ranging from 0-4. The overall judgment of level of activity should override any specific behavioral examples.
0 = Extremely low activity: Lying on the quilt with very little movement except looking around or quietly crying. 1 = Low activity: Partial body movements: somewhat more active. The child may "wiggle" its arms. 2= Moderate activity: More active movement of arms and legs or some whole-body movement on the quilt. 3 = High activity: Very active movement of arms and legs or more active whole-body movements. 4 = Extremely high activity: Thrashing of arms and legs and the related whole-body movements.

*__Note__: Because the age of babies participating in Prelocomotor Lab-TAB, it is important to remember that some babies may be crawling while others may not. Therefore, intensity of movement must be viewed strictly as intensity of any movement rather than a certain type of movement being associated as indicating higher or lower activity.

d) Parent behavior: 0= Interfering; emotionally loaded statements to C, soothing, reprimanding C, commanding, or generally disrupting. 1= Mild interference; 1-2 comments directed at C or adjustments of C. These comments or adjustments are not emotionally loaded. 2= Not interfering, neutral.

e) Baseline state: The child's state prior to the beginning of the episode:1= tired/drowsy. 2= alert/calm. 3= alert/active. 4= fussy. 5= crying.

10. EC 2,4. RESTRAINT (Anger/Frustration)

Rationale

Being physically restrained or compelled to do something against one's wishes can elicit anger. Restraint during the length measurement of the anthropometry assessment is intended to elicit mild anger responses in some children.

Scoring

The episode begins when E restrains C on the board and is divided into three 5-second epochs (total = 15 seconds) if there is enough time given to this. The epochs are coded by indicating the occurrence of the specified behavior, or by rating the intensity of the behavior. When an intensity rating is requested, the highest intensity observed should be coded.

Variables to be coded: a) Latency to anger response. b) Latency to sadness response. c) Intensity of facial anger. d) Intensity of facial sadness. e) Intensity of distress vocalizations. f) Intensity of struggle. g) Baseline state. h) Parent behavior.

Definitions of variables:

a) Latency to anger response: Interval, in seconds, to the first sign of anger after C is put on the board. Latency to anger should only be recorded if c) Intensity of facial anger is coded at 1 or greater. No other indicators of anger should be measured here (e.g., crying) unless an anger face is present.

b) Latency to sadness response: interval, in seconds, to the first sign of sadness (facial, postural, or vocalic). Latency to sadness should only be recorded if d) Intensity of facial sadness is coded at 1 or greater. No other indicators of sadness should be measured here (e.g., crying) unless a sad face is present. Because of the way that restraint was carried out, Presence of bodily sadness will not occur. (This variable has been deleted from this list.)

c) Intensity of facial anger: Peak intensity of facial anger or anger blends is noted in each epoch using AFFEX (See Appendix A for definitions) and rated on the following scale: 0= No facial region shows codable anger movement. 1= Only one facial region shows codable movement, identifying a low intensity anger, or expression is ambiguous. 2= Only 2 facial regions show codable anger movement or movement is very clear in one facial region. 3= An appearance change occurs in all 3 facial regions, or coder otherwise has impression of strong anger.

d) Intensity of facial sadness: Peak intensity of facial sadness or sadness blends is noted in each epoch using AFFEX (See Appendix A for definitions) and rated on the following scale: 0= No facial region shows codable sadness movement. 1= Only one facial region shows codable movement; identifying very low or ambiguous facial

sadness. 2= Only 2 facial regions show codable movement, or expression in one region is very clear. 3= An appearance change occurs in all 3 facial regions or coder otherwise has impression of strong facial sadness.

e) Intensity of distress vocalizations: Peak intensity of distress vocalizations is noted in each epoch and rated on the following scale: 0= No distress. 1= Mild protest verbalization that may be difficult to identify as hedonically negative. 2= Definite protest, limited to a short (1-2 second) duration. 3= Longer protest, fussing or mild, low-intensity cry (cry has extended or rhythmic quality). 4= Definite non-muted crying. 5= Full intensity cry/scream (child is losing control).

***Note**: Some vocalizations in this episode will not be anger-related and should not be coded.

f) Intensity of struggle: Peak intensity of struggling in each epoch is rated on the following scale: 0= No struggling at all. No resistant movement. 1= Low intensity struggle. Examples include 1-2 instances of pushing against the restraints, shifting, twitching, light wiggling, and low intensity kicking. 2= Medium intensity struggle. Full body movements such as 2 or more intense pushes against the car seat restraints, twisting, leaning forward, arching back or kicking.
3= Moderately high intensity struggle. Near continuous moderate intensity movements including those from '2' with higher intensity. These movements usually last more than half of the epoch.
4= High intensity struggle. Continuous movement of moderately high intensity with intervals of high intensity resistance. Struggle lasts the entire epoch.

g) Baseline state: The child's state prior to the beginning of an episode: 1= tire/drowsy. 2= alert/calm. 3= alert/active. 4= fussy. 5= crying.

h) Parent behavior 0= Interfering; emotionally loaded statements to C, soothing, reprimanding C, commanding, or generally disrupting. 1= Mild interference; 1-2 comments directed at C or adjustments of C. These comments or adjustments are not emotionally loaded. 2= Not interfering: neutral.

APPENDIX A: AFFEX FACIAL EXPRESSION DEFINITIONS

Emotion	Movements In Forehead/Brows Regions	Movements In Eyes/Nose/Cheeks Regions	Movement In Mouth/Lips/Chin Regions	Notes
Anger	V Inner corners are lowered and drawn together. V Bulging or vertical furrows between the eyes may be visible due to this movement.	V Eyes may look tense or squinted. V Cheeks may be raised. V Fold under eye may deepen.	V Mouth looks tense, wide open and squarish. V Alternatively, mouth appears closed with lips pressed together.	V Don't confuse brow movements with those in V interest. See illusion of sadness note.
Fear	V Entire brow should be raised/neutral and drawn together. V Brows may also look straighter across than usual. V Faint horizontal furrows may be present in forehead.	V Upper eyelid raises making the eyes appear wider. V Eyes have tense appearance.	V Lip corners are drawn straight back. V Mouth is usually less than wide open.	V Don't confuse interest brows V for fear. See illusion of sadness note.
Sadness	V Inner corners move upward and together resulting in bulging/furrows in middle of forehead.	V Cheeks may look lower than usual or have a droopy appearance. V Alternatively, cheeks may be raised and eyes squinted.	V Lip corners should be drawn down. V Bottom lip may be pushed up and out by the chin which may be tense or wrinkled.	
Joy	V Most likely remain neutral.	V Cheeks raise V Furrow below the eyes deepens. V "Crows feet" will extend from the outer corners of the eye. V Eyes may appear squinted.	V Lip corners are raised. V Nasolabial fold deepens.	
Interest	V Entire brow is raised. V Alternatively, brows are drawn together and slightly lowered.	V Eyes look wider than usual due to raised brows. V Alternatively, eyes may be squinted and cheeks raised.	V Mouth may open.	V When coding infants, do not code "mouth opens" as interest unless it is in response to a stimulus

Note on the potential for an "illusion of sadness"

There are several occasions when an illusion of sadness may appear. Sadness should not be coded in these situations:

V The first situation is when brows are drawn tightly down and together. In this case, it is common for the inner most corners of the brows to bulge up in the middle falsely giving the appearance of sadness. This is most likely due to the large amount of fat in the infant face.

V The second situation is when the outer corners of the brows are lowered falsely giving the appearance that the inner corners have raised. In this case, be sure to observe the actual movement of the brows. In sadness, the inner corners need to be raised and drawn together. Simply observing a still frame of this expression is not sufficient to distinguish between true sadness and the illusion of sadness.

V Finally, an illusion of sadness may occur when children inhale deeply during a bout of crying. In this situation, the lip corners will be drawn down by the inhaling action giving the impression of sadness.

These descriptions were adapted from C.E. Izard's *The Maximally Discriminative Facial Movement Coding System.*

APPENDIX B: CODING SHEETS

Shortened Version of the Ethiopian 6-month LabTab Study
Spring 2008

1. Puppet Game EC:3,2

Subject number_____	Scorers:
Date scored _____	

Latency to pleasure response_____s
Latency to approach puppets in epoch 5_____s

Epochs	1	2	3	Avg.
Time (Begin/End)				
Intensity of Smiling (0-3)				
Presence of Laughter (0-1)				
Presence of Positive Vocalizations (0-1)				
Presence of Positive Motor Acts (0-1)				

If all zeros: _____negative or _____neutral reaction

Engaged With Toy:_____ Missing episode code _____

Parent behavior:_____ # of observed epochs _____

Baseline state: _____

Shortened Version of the Ethiopian 6-month LabTab Study
Spring 2008

2. EC 1,1. PARASOL OPENING (Fear)

Subject number _____	Scorers:
Date scored _____	

Latency to fear response T1 _____s T2 _____s T3 _____s

	1	2	3	Avg.
Time (Begin/End)				
Intensity of Facial Fear (0-3)				
Intensity of Vocal Distress (0-5)				
Intensity of Bodily Response (0-3)				
Intensity of Escape (0-3)				
Intensity of Startle (0-3)				

Parent behavior:_____

Baseline state:_____

Missing episode code _____

# of observed epochs _____

Shortened Version of the Ethiopian 6-month LabTab Study
Spring 2008

3. EC 3,4. MODIFIED PEEK-A-BOO GAME (Joy/Pleasure)

Subject number_____	Scorers:
Date scored _____	

Latency to pleasure response _____ s

	Epochs			
	1	2	3	Avg.
Time (Begin/End)				
Intensity of Smiling (0-3)				
Presence of Laughter (0-1)				
Presence of Positive Vocalizations (0-1)				
Presence of Positive Motor Acts (0-1)				

Parent Performance_____
Missing episode code _____

Baseline state _____
# of observed epochs _____

If all zeros: _____negative or _____neutral reaction

Shortened Version of the Ethiopian 6-month LabTab Study
Spring 2008

4. EC TASK ORIENTATION (BLOCKS) (Interest/Persistence)

Subject number_____	Scorers:
Date scored _____	

Latency to look away from blocks _____ s

Epochs

	1	2	3	4	5	6	7	8	9	Avg.
Time (Begin/End)										
Duration of Looking (0-3)										
Manipulation of Stimuli (0-3)										
Intensity of facial interest (0-2)										

Parent behavior_____
Missing episode code _____

Baseline state_____
# of observed epochs _____

Shortened Version of the Ethiopian 6-month LabTab Study
Spring 2008

5. EC 2,3. TOY RETRACTION (Anger/Frustration)

Subject number_____	Scorers:
Date scored _____	

Latency to any anger: T1 _____s T2 _____s

Latency to any sadness: T1 _____s T2 _____s

	Trial 1			Trial 2			Avg trial 1.	Avg. trial 2
5 sec. epochs	1	2	3	1	2	3		
Time (Begin/End)								
Intensity of Struggle (0-4)								
Intensity of Facial Anger (0-3)								
Intensity of Distress Vocalizations (0-5)								
Intensity of Facial Sadness (0-3)								

Maximum Struggle T1 _____ T2 _____ T3 _____

Interest In Toy_____

Baseline state _____ Parent behavior_____

Missing Episode Code_____ # of observed epochs_____

Shortened Version of the Ethiopian 6-month LabTab Study
Spring 2008

6. EC 5,3. BASKET OF TOYS (Activity Level)

Subject number_____	Scorers:
Date scored _____	

Latency to approach any toy _____s

20 sec. epochs	1	2	3	4	5	6	Avg.
Time (Begin/End)							
Intensity of manipulation (0-4)							
Bouts of toy play							

Parent behavior:_____
Missing Episode Code_____

Baseline state:_____
# of observed epochs_____

Shortened Version of the Ethiopian 6-month LabTab Study
Spring 2008

7. EC 4,2. PERSON INTEREST (Interest/Persistence)

Subject number_____	Scorers:
Date scored _____	

Latency to look away from person_____ s

Epochs

10 s epochs	1	2	3	4	5	6	Avg.
	Talk	Talk	Talk	Stand	chair	chair	
Time (Begin/End)							
Duration of Looking (0-3)							
Verbal Comments to/about person (1,0)							
Intensity facial interest (0-2)							

Parent behavior_____
Missing episode code _____

Baseline state _____
# of observed epochs _____

Shortened Version of the Ethiopian 6-month LabTab Study Spring 2008

8. EC 1,3. STRANGER APPROACH (Fear)

Subject number_____	Scorers:
Date scored _____	

Latency To Fear Response: S1_____ S2_____ S3_____ Latency To Sadness Response: S1_____ S2_____ S3_____

	Stage 1				Stage 2			Stage 3		
	Stranger Enters	Moves Toward C	Pauses and Speaks		Moves Near C	Pauses Near C		Picks up Holds C	Replaces C	
Time Begin/End				Avg.			Avg.			Avg.
Intensity of facial fear (0-3)										
Intensity of facial sadness (0-3)										
Intensity of vocal distress (0-5)										
Intensity of bodily fear (0-3)										

Presence of bodily sadness 0=no 1=yes										
Intensity of escape behavior (0-3)										

*approx. Baseline state _____ Missing Episode Code _____

Parent behavior_____ # of observed epochs _____

Shortened Version of the Ethiopian 6-month LabTab Study
Spring 2008

9. EC 5,2. PRONE & SUPINE PLACEMENT (Activity Level)

Subject number_____	Scorers:
Date scored _____	

Latency to move:
T1_____ T2_____

Latency to turn over:
T1_____ T2_____

Trial 1 (prone) Trial 2 (supine)

	1	2	Avg.	3	4	Avg.
Time begin/end						
Intensity of movement (0-4)						

Parent behavior_____

Missing Episode Code_____

Baseline state_____

# of observed epochs_____

Shortened Version of the Ethiopian 6-month LabTab Study
Spring 2008

10. EC 2,4. RESTRAINT

Subject number_____	Scorers:
Date scored _____	

Latency to anger response _____s Latency to sadness response _____s

5 s epochs	1	2	3	Avg.
Time (Begin/End)				
Intensity of Facial Anger (0-3)				
Intensity of Facial Sadness (0-3)				
Intensity of Distress Vocalizations (0-5)				
Intensity of Struggle (0-4)				

Parent behavior _____ Missing Episode Code _____

Baseline state _____ # of observed epochs _____

APPENDIX C: EMOTION REGULATION CODING

I. Duration of attention: amount of time child is looking at the stimulus, scored on an intensity scale (or could use real time)

II. Disengagement of attention Gaze aversion: child looks away from stimulus without focusing on any particular object - this behavior is extremely brief in duration (score as present or absent)

Distraction toward object: child looks at an object that is unrelated to the episode - this is usually for a longer duration than gaze aversion (score as present or absent)

III. Approach/Withdrawal: child approaches or withdraws self from stimulus (can be scored as present or absent, or on an intensity scale). When using an intensity scale, we recommend using approach and withdrawal as separate scales.

IV. Social Strategies Looks to mother: child looks to mother - it is possible to distinguish between the types of looks, for example, positive and negative affect or social referencing and information seeking (score as present or absent)

Looks to experimenter: child looks to experimenter (score the same as looks to mother

V. Dealing with the stimulus (these behaviors can be combined into a composite) **Exploring**: child not only attends to the stimulus but inspects it with concentration in an attempt to understand how it works (score as present or absent)

Struggling/resisting: child pulls, kicks, arches his/her back, pushes etc. (score as present or absent)

Control: child controls situation by attempting to move stimulus, (e.g., push it away) (score as present or absent)

Playing: child plays with stimulus in an appropriate manner (score as present or absent)

VI. Redirected action Self-stimulation: child uses a body part to engage in repetitive manipulation (e.g., sucking thumb) (score as present or absent)

Tension release: child engages in high-energy behavior with no apparent instrumental focus (e.g., screaming or fast kicking of the legs) (score as present or absent)

APPENDIX C

Institutional Review Board Approval

Figure 3

Institutional Review Board Approval

Oklahoma State University Institutional Review Board

Date:	Wednesday, March 19, 2008
IRB Application No	AS0826
Proposal Title:	A Validation of an Ethiopian Version of the Laboratory Assessment Battery
Reviewed and Processed as:	Exempt

Status Recommended by Reviewer(s): Approved Protocol Expires: 3/18/2009

Principal Investigator(s):

Stephanie Grant	David Thomas
1000 Hyacinth Hollow Drove	116 N. Murray
Yukon, OK 73099	Stillwater, OK 74078

The IRB application referenced above has been approved. It is the judgment of the reviewers that the rights and welfare of individuals who may be asked to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

▣ The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

As Principal Investigator, it is your responsibility to do the following:

1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval.
2. Submit a request for continuation if the study extends beyond the approval period of one calendar year. This continuation must receive IRB review and approval before the research can continue.
3. Report any adverse events to the IRB Chair promptly. Adverse events are those which are unanticipated and impact the subjects during the course of this research; and
4. Notify the IRB office in writing when your research project is complete.

Please note that approved protocols are subject to monitoring by the IRB and that the IRB office has the authority to inspect research records associated with this protocol at any time. If you have questions about the IRB procedures or need any assistance from the Board, please contact Beth McTernan in 219 Cordell North (phone: 405-744-5700, beth.mcternan@okstate.edu).

Sincerely,

Shelia Kennison, Chair
Institutional Review Board

VITA

Stephanie Lea Grant

Candidate for the Degree of

Master of Science

Thesis: A VALIDATION OF AN ETHIOPIAN VERSION OF THE LABORATORY TEMPERAMENT ASSESSMENT BATTERY

Major Field: Psychology

Biographical:

- **Personal Data:**
  (405) 248-8633
  1000 Hyacinth Hollow Drive, Yukon, OK 73099

- **Education:**
  - Completed the requirements for the Master of Science in Psychology at Oklahoma State University, Stillwater, Oklahoma in December, 2008.
  - Completed the requirements for the Master of Arts in Marriage and Family Therapy at Southern Nazarene University, Bethany, Oklahoma in May, 2006.
  - Completed the requirements for the Bachelor of Science in Psychology at Southern Nazarene University, Bethany, Oklahoma in May, 2004.

- **Experience:**
  - Assistant Professor of Psychology, June 2006 – Present
    Southern Nazarene University, Bethany, Oklahoma
  - Coordinator of the Student Advising Center, July 2005 – June 2006
    Oklahoma Wesleyan University, Bartlesville, Oklahoma
  - Counseling Intern, October 2005 – May 2006
    Wilson Psychological Associates, Bartlesville, Oklahoma

- **Professional Memberships:**
  - American Psychological Society, January 2007 to Present
  - Southwestern Psychological Association, October 2003 to Present
  - Oklahoma Psychological Society, March 2008 to Present

Name: Stephanie Lea Grant Date of Degree: December, 2008

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: A VALIDATION OF AN ETHIOPIAN VERSION OF THE LABORATORY TEMPERAMENT ASSESSMENT BATTERY

Pages in Study: 107 Candidate for the Degree of Master of Science

Major Field: Psychology

Scope and Method of Study:

Temperament is a well-researched construct that is often of interest when attempting to account for individual differences in data. Several models explaining temperament exist and studies of temperament that focus on early infancy in particular, like the current study aimed to do, remain of high interest because they are able to capture temperamental behaviors before the environment is able to play a large role in the developmental process. While temperament has found its way into an abundance of research, the assessment of temperament, particularly cross-culturally, seems to lack consistency and reliability. Continued cross-cultural study is yet very relevant due to the increasing amount of literature suggesting the developmental importance of the construct. The purpose of the present study was thus to provide evidence of the validity of an Ethiopian version of an infant temperament assessment using the Laboratory Assessment Battery (Lab-TAB), as well as validity for a shortened version of this cross-cultural measure. This was done by examining the psychometric properties of the instrument with a sample of 6-month Ethiopian infants.

Findings and Conclusions:

Results from the present study do suggest continued use of Lab-TAB with an Ethiopian population as well as support for the shortened version of this measure. As the first study to assess the use of Lab-TAB in cross-cultural research and with the developmental importance of temperament, this work makes a large contribution to the cross-cultural temperament literature available.

ADVISER'S APPROVAL: David G. Thomas