HHS Public Access

Author manuscript

J Pers Assess. Author manuscript; available in PMC 2022 March 01.

Published in final edited form as:

J Pers Assess. 2021 ; 103(2): 214–223. doi:10.1080/00223891.2020.1722140.

Using a Thin Slice Coding Approach to Assess Preschool Personality Dimensions

Diana J. Whalen, PhD[1], **Kirsten E. Gilbert, PhD**[1], **Joshua J. Jackson, PhD**[3], **Deanna M. Barch, PhD**[1,2,3,4], **Joan L. Luby, MD**[1]

[1]Department of Psychiatry, Washington University in St. Louis

[2]The Program in Neuroscience, Washington University in St. Louis

[3]Department of Psychological and Brain Sciences, Washington University in St. Louis

[4]Department of Radiology, Washington University in St. Louis

Abstract

Background: A large literature assessing personality across the lifespan has used the Big Five as an organizing framework, with evidence that variation along different dimensions predicts aspects of psychopathology. Parent reports indicate that these dimensions emerge as early as preschool, but there is a need for objective, observational measures of personality in young children, as parent report can be confounded by the parents' own personality and psychopathology.

Methods: The current study observationally coded personality dimensions in a clinically enriched sample of preschoolers. A heterogeneous group of preschoolers oversampled for depression (N=299) completed 1-8 structured observational tasks with an experimenter. Big Five personality dimensions of extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience were coded using a "thin slice" technique with 7,820 unique ratings available for analysis.

Results: Thin-slice ratings of personality dimensions were reliably observed in preschoolers ages 3-6 years. Within and across-task consistency was also evident, with consistency estimates higher than found in adult samples. Divergent validity was limited, with coders distinguishing between three (extraversion/openness; agreeableness/conscientiousness; neuroticism) rather than five dimensions.

Conclusions: Personality dimensions can be observationally identified in preschool-age children and offer reliable estimates that stand across different observational tasks. Study findings highlight

Address Correspondence to: Diana J. Whalen, PhD, Washington University School of Medicine, Department of Psychiatry, 4444 Forest Park, Suite 2100, St. Louis, MO 63108, Phone: 314-286-2730; Fax: 314-286-2732; diana.whalen@wustl.edu.

DJW- study conceptualization, data collection, data preparation, data analysis, report writing
KEG- study conceptualization, data collection, data preparation, data analysis, report writing
JJJ- data preparation, data analysis, report writing
DMB- study conceptualization, data collection, report writing
JLL- study conceptualization, data collection, report writing

DECLARATION OF CONFLICTING INTERESTS
The authors declare no conflicts of interest.

the importance of observational approaches to assessing early personality dimensions, as well as the utility of the thin slice approach for meaningful secondary data analysis.

Keywords

personality development; thin-slice coding; preschool personality

Introduction

The Big Five, or Five Factor Model (FFM) is an organizing framework at the heart of our understanding of personality traits across the lifespan. DSM-5 (American Psychiatric Association, 2013) has put forth an alternative model of personality disorders based upon, and overlapping considerably with, the Big Five (Markon, Krueger, & Watson, 2005), highlighting the importance of this framework in understanding links between personality dimensions and psychiatric illness. The Big Five measures personality along five dimensions: extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience (Caspi, Roberts, & Shiner, 2005; Goldberg, 1993). Although individual differences in Big Five dimensions have been measured in children as young as preschool, this has been done almost exclusively through parent reports (De Pauw, 2017; Herzhoff, Kushner, & Tackett, 2017). In contrast, the current study used a thin slice coding approach to examine whether Big Five dimensions could be reliably identified from observations of preschoolers by unacquainted observers.

Personality in preschoolers

Individual differences in the thinking, feeling, and behavior patterns of preschoolers have been traditionally studied under the broader concept of temperament, as opposed to personality, due to the belief that early temperament may be more biologically based than adult personality traits. However, recent work challenging these conceptual distinctions between temperament and personality (De Pauw, 2017; De Pauw & Mervielde, 2010; McCrae et al., 2000; Shiner, Masten, & Roberts, 2003; Victor, Rothbart, Baker, & Tackett, 2016) has led to the development of models of personality extended to preschoolers (Caspi et al., 2005; Shiner et al., 2003; Tackett et al., 2012). Many of the Big Five dimensions have clear complements in the temperament literature, for example: extraversion-surgency/sociability, neuroticism-negative emotionality, and conscientiousness-effortful control. Thus, research on preschool and early childhood personality dimensions has expanded considerably over the past decade (Abe, 2005; Abe & Izard, 1999; Asendorpf & Denissen, 2006; De Pauw, Mervielde, & Leeuwen, 2009; Goldstein et al., 2019; Grist & McCord, 2010; Halverson et al., 2003; Lamb, Chuang, Wessels, Broberg, & Hwang, 2002; Measelle, John, Ablow, Cowan, & Cowan, 2005; Olino, Klein, Dyson, Rose, & Durbin, 2010; Soto, 2016; Tackett et al., 2012; Vollrath, Hampson, & Torgersen, 2016; Wilson, Schalet, Hicks, & Zucker, 2013; Zupančič, Podlesek, & Kavčič, 2006). Collectively, this work indicates that individual differences in preschool-age children can be reliably described using personality dimensions, such as the five-factor model, and that preschool personality dimensions can predict longitudinal outcomes in several domains including social, health, neural, and psychological functioning.

In most recent studies, preschool personality dimensions have been measured in two ways: parent-report and observational protocols. Given the young age of the children, self-reports of personality have been assumed to be unreliable and rarely used (see Measelle et al., 2005 for an exception), thus primary caregivers have been relied upon to give accurate and valid information on their preschooler's personality. Parent-report methods have known limitations including informant bias and role or context specificity (Durbin & Wilson, 2012). The temperamental counterparts of personality traits in young children have also been assessed via observational protocols, such as the Laboratory Temperament Assessment Battery, which tend to be resource intensive, taking large administrative and coding efforts (Dyson et al., 2015; Dyson, Olino, Durbin, Goldsmith, & Klein, 2012; Kopala-Sibley, Olino, Durbin, Dyson, & Klein, 2018; Kryski et al., 2018). Although shown to be reliable and valid in older samples, no previous study has utilized a thin slice coding technique for the assessment of personality dimensions in preschoolers. Thin slice coding techniques can offer significant advantages when compared to the coding techniques typically used to study personality dimensions. First, thin slice coding is significantly more efficient and less resource intensive than traditional coding. Limited time is spent "training" coders and coders are not required to reach a reliability benchmark prior to coding. Thin slice coding is primarily based off of the coder's perception of a child, rather than a formal manual with heavily detailed descriptions of codes, levels, and/or anchors that must be strictly followed in order to maintain reliability. As such, thin-slice coding takes significantly less time than traditional coding protocols. Another advantage of the thin slice coding technique is that it can be used on almost any available video-taped (or audio-taped) interactions with a child, making this approach ideal for secondary data analysis.

Thin Slice assessments of personality

Personality is manifest in everyday behaviors that are visible to friends, family, and close others, thus allowing observers to accurately rate another's personality (Vazire & Carlson, 2011). One does not have to know someone to accurately rate personality, as naïve observers who have never met a person can provide accurate ratings based on brief periods of observed behavior (Jackson et al., 2009). This ability extends to short, "thin-slice" assessments whereby a naïve observer views approximately 30-second to one-minute videos of a person in different contexts. In adult samples, thin-slice assessments by observers: (1) show agreement and consistency with other observers, (2) show agreement with self-reports of personality by the target individual, and (3) demonstrate consistency across situations (Ambady & Rosenthal, 1992; Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004; Borkenau & Zaltauskas, 2009; Carcone et al., 2015; Murphy et al., 2019, 2015; Slepian, Bogart, & Ambady, 2014). Together, these findings highlight an untrained individual's ability to quickly and accurately process information relevant to personality ratings.

Thin slice approaches are almost exclusively used with adult samples. However, four recent studies incorporated thin slice ratings of childhood and early adolescent personality (Prime, Perlman, Tackett, & Jenkins, 2014; Tackett, Herzhoff, Kushner, & Rule, 2016; Tackett, Lang, Markon, & Herzhoff, 2019; Tackett et al., 2017), offering initial evidence of validity and high cost-efficiency. In each of these studies, children's personality dimensions were reliability rated by unacquainted observers who watched brief behavioral observations of

each child. So far, this method has only been applied to older children and it is unclear whether thin slice observational ratings can adequately assess personality dimensions during the preschool (ages 3-6) period, a key developmental stage increasingly recognized as initiating pathways for lifelong health and well-being. Further, little is known about the cross-situational consistency in thin slice ratings during this stage of development. It is important to investigate whether naïve observers can consistently rate a young child's personality based on observation of behavior across unique situations as doing so will inform what situations are ideal for rating certain traits, as well as identify situations that may constrain personality manifestations, which would limit their utility in creating valid personality assessments.

The current study addressed these issues by using a thin slice coding approach to assess Big Five personality dimensions in a large heterogeneous preschool sample enriched for clinical psychopathology. The main objective was to investigate whether Big Five personality dimensions could be reliably identified in preschoolers and to look at cross-situational consistency in ratings of personality dimensions. There are a limited number of studies that have used this approach with children, yet those studies offer promising reliability and validity estimates. Therefore, we hypothesized that each of the preschool personality dimensions would be able to be reliably discerned using this technique. Convergent and divergent validity were also explored, with no explicit predictions made in advance.

METHODS

Participants

Participants included 305 young children enrolled in the Preschool Depression Study (PDS; Luby, Belden, Pautsch, Si, & Spitznagel, 2009; Luby, Si, Belden, Tandon, & Spitznagel, 2009). The PDS is a prospective longitudinal study of preschool-age children designed to investigate the validity and longitudinal course of preschool depression, conducted at the Washington University School of Medicine Early Emotional Development Program in St. Louis. For the PDS, 3.0- to 5.92 year-old children and their primary caregivers were recruited from daycares, preschools, and primary care sites, using the Preschool Feelings Checklist (Luby, Heffelfinger, Koenig-McNaught, Brown, & Spitznagel, 2004) to oversample children with depression or at-risk for depression. Healthy preschoolers and those with other psychiatric disorders were also included in the study via similar recruitment methods. Children have since undergone annual diagnostic and developmental assessments (i.e., approximately every 12 months). Parental consent and child verbal assent were obtained before study participation. The Institutional Review Board approved all procedures in accordance with institutional ethical guidelines.

Procedures

As part of the larger PDS study, annual in-person assessments were conducted beginning during the preschool period (Luby, Si, et al., 2009). Approximately 65% (n= 197) of children were diagnosed with a psychiatric disorder at baseline (ages 3-6 years). Children completed tasks from the LABTAB assessment battery (Goldsmith, Reilly, Lemery, Longley, & Prescott, 1999) during the preschool period in this ongoing longitudinal study (see Table 1

for list of tasks). For the current report, 8-18 unique observers rated each child's personality dimensions (e.g., Big Five) using video previously recorded during each of the LABTAB tasks conducted as part of the larger project.

Measures

Income to needs.—Mothers reported on the family income and number of individuals in the household supported by this income between the ages of 3-6 years. An income-to-needs ratio was computed as the total family income divided by the federal poverty level, based on family size, at the time of data collection (McLoyd, 1998). An income-to-needs ratio less than one is indicative of low-income/poverty status. Income-to-needs was used as a control variable in our analyses given the relatively large percentage of low-income families in our sample as well as prior work indicating that raters of adult samples using the thin slice technique are able to reliably determine socio-economic status of participants (Kraus & Keltner, 2009).

Personality Dimensions.—To obtain the best approximation of preschool personality dimensions, observational data was coded using a thin slice technique and aggregated from the first three annual assessments (representing the preschool-age period). Thus, 7,820 unique ratings of children during ages 3-6 years formed the basis for the thin slice assessments by unacquainted observers. The unacquainted observers were 27 undergraduate research assistants and Bachelors/Masters-level staff of the Early Emotional Development Program. Observers varied in terms of education level, race, ethnicity, gender, and age. Each observer was oriented to the thin slice protocol and then completed his/her ratings of each child. The orientation session briefly described the thin slice procedure and reviewed each rating that was to be made (Table 2). Observers remained blind to child diagnostic and social characteristics (e.g., income-to-needs) that may have unintentionally influenced ratings. At least 8 and up to 18 unique observers rated each child's personality dimensions (e.g., Big Five) during each of the LABTAB tasks conducted as part of the larger PDS study. These tasks included a standard set of instructions given by a research assistant, who interacted with the child apart from his/her caregiver.

In order to make personality ratings, each observer watched and listened to approximately 60-seconds (taken from the middle) of each structured observational task as done in prior work using the thin slice technique with adult samples (Carney, Colvin, & Hall, 2007; Hirschmann, Kastner-Koller, Deimann, Schmelzer, & Pietschnig, 2018; Murphy et al., 2015). The middle portion of each task was preselected and cut out of the remaining videotaped task to decrease potential for bias (e.g., watching portions before or after the selected 60-second clip) and to replicate existing studies using this technique in adult samples (Murphy et al., 2019, 2015). Each video was coded for all Big Five (extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience) personality dimensions using a 5-point likert scale (e.g., 1- not at all, 2- a little, 3-moderately, 4-quite a bit, 5-extremely). For example, after watching the child for 60 seconds, observers were asked to rate how "extraverted" the child was using the likert scale. Coders were given brief definitions of each personality dimension (e.g., "extraverted: talkative, assertive, active, excitement-seeking, and fun loving"). Each description given to coders is provided in Table

2. Ratings were averaged across observers for each personality dimension with an average of 25.7 (SD=5.13; Min=10, Max=33) unique ratings included for each child. Six children with less than 10 ratings were excluded, resulting in a final sample of n=299.

Temperament.—In order to examine convergent validity with thin slice ratings, parents completed a modified version of the Children's Behavior Questionnaire (Rothbart, Ahadi, Hershey, & Fisher, 2001) during the preschool behavioral assessment. The Children's Behavior questionnaire (CBQ) is a widely used, reliable assessment of temperament dimensions during early childhood, such as negative affectivity, surgency/extraversion, and effortful control. For the larger study, only specific subscales were administered to caregivers. These included: anger/frustration, approach, falling reactivity/soothability, fear, high intensity pleasure, low intensity pleasure, sadness, shyness, and smiling/laughter.

Statistical Analysis

Independent samples t-tests and Pearson correlations were used to calculate descriptive statistics on the personality dimensions and how these related to age, sex, and family income-to-needs ratio. Intraclass correlation coefficients (ICCs) were calculated for inter-rater agreement and cross-situational consistency for each personality dimension. Convergent and divergent validity were calculated with partial correlations.

RESULTS

Descriptives for Personality Dimensions

Descriptive statistics and correlations for demographic variables and thin-slice ratings of personality dimensions are reported in Table 3. There were 144 female participants and 155 male participants. Several sex differences emerged in thin slice ratings of preschool personality dimensions. Girls (M= 3.18, SD= 0.45) were rated as being more conscientious (t_{297}= −3.65, p<0.001; d=.41) and agreeable (M= 3.46, SD= 0.40; t_{297}= −3.26, p<0.001; d= .39) than boys (M= 3.00, SD= 0.39; M= 3.32, SD= 0.38). Boys (M= 2.19, SD= 0.28) were rated as being higher on neuroticism than girls (M= 2.07, SD= 0.28; t_{297}= 3.67, p<0.001; d=.43). There were no sex differences in ratings of extraversion (t_{297}= −0.56, p=0.96; d=.02) or openness (t_{297}= 0.91, p=0.36; d=.09).

Bivariate correlations were conducted to determine if preschool personality dimensions were associated with the child's age and family income-to-needs ratio. Child age was positively associated with conscientiousness (r= 0.33, p<0.001) and agreeableness (r= 0.25, p<0.001) and negatively associated with openness (r= −0.19, p<0.001). No age differences were noted in the ratings of extraversion (r= −0.10, p=0.08) and neuroticism (r= −0.09, p=0.14). Family income-to-needs ratio was positively associated with higher ratings on conscientiousness (r= 0.37, p<0.001) and agreeableness (r= 0.22, p<0.001). No family income-to-needs differences were noted in the ratings of extraversion (r= −0.10, p=0.11), openness (r= −0.01, p=0.81), or neuroticism (r= 0.06, p=0.33).

Inter-rater Agreement

Intraclass correlation coefficients (ICCs) among raters' scores of personality dimensions are displayed in Tables 4 and 5. We used the same types of ICCs reported in past work using the thin slice coding technique as a template for the current study (Borkenau et al., 2004; Koo & Li, 2016; Tackett et al., 2016). We first examined the average agreement between individual raters who observed the same task using one-way random effects, absolute agreement ICC (1,1; single rater/ single task) and the average agreement of the composite score across tasks of the range of raters who observed the same task using one-way random effects, absolute agreement ICC (1, *k*; average rater/single task). We also examined the average cross-task/ cross-rater correlation between ratings using two-way random effects, absolute agreement ICC (2,1; mean rater, 8 tasks).

Agreement between individual raters who observed the same task (e.g., the ICC using single ratings for a single task) ranged from 0.24 (openness) to 0.53 (extraversion) with a mean of 0.33. Two random raters' personality assessment for a single task would correlate roughly .3, which is similar in magnitude to the agreement levels seen using the thin slice technique in adults (Connelly & Ones, 2010). Given that we had 8-18 unique raters for each child, we also estimated the reliability equivalent to Chronbach's alpha for *k* raters assessing the same thin slice (e.g., the ICC using the averaged ratings for a single task). Agreement of all the raters who observed the same task (average of 3) ranged from 0.46 (openness) to 0.77 (extraversion) with a mean of 0.57. Reliability was also estimated by collapsing across tasks to estimate agreement for an overall personality dimension rating (e.g., the ICC using the averaged ratings across tasks). ICCs were similar in magnitude compared to the estimates for a single task, with ICCs ranged from 0.37 (neuroticism) to 0.67 (extraversion) with a mean of 0.47. This range of ICCs is equivalent to prior research using the thin slice technique to assess older children's personality dimensions (Tackett et al., 2016, 2017).

Across all personality dimensions, the *Storytelling* task had the highest average rater/ single task ICCs (ranging from 0.52 to 0.85), whereas the *Empty Box* task had the lowest average rater/ single task ICCs (ranging from 0.11 to 0.68). For individual personality dimensions the highest average rater/ single task ICC for extraversion (0.85) was from the *Storytelling* task, for agreeableness (0.69) was from the *Snack Delay* task, for conscientiousness (0.77) was from the *Snack Delay* task, for neuroticism (0.64) was from the *Storytelling* task, and for openness (0.65) was from the *Storytelling* task.

Cross-situational Consistency

It is important to examine rating consistency across contexts as consistency implies the existence of stable characteristics being assessed (e.g., personality traits) rather than raters being able to view and agree on state relevant behaviors. While ICCs averaged across all tasks speak to this in some degree, the cross-situational consistency of children's thin slice personality dimension ratings directly addresses this issue (e.g., cross-rater/cross-task ICC divided by average rater/single task ICC; Table 4; (Borkenau et al., 2004)). See Supplemental Tables 1–5 for correlations of personality dimensions across tasks. Cross-situational consistency estimates ranged from 0.76 (neuroticism) to 0.88 (extraversion) with a mean of 0.82. These estimates are notably higher than those derived in previous work

adults (Borkenau et al., 2004; Borkenau & Zaltauskas, 2009) and consistent (albeit slightly higher) with previous estimates using slightly older children and adolescents (Tackett et al., 2016, 2017). Our cross-task consistency estimates were highest for extraversion, which matches with past work in adults and older children (Borkenau et al., 2004; Tackett et al., 2016). However, our cross-task consistency estimates were lowest for neuroticism, which has not been found in older children and adults.

Exploratory analysis of convergent and divergent validity

Thin slice ratings were found to be reliably assessed within a task, for multiple different types of tasks, and collapsed across tasks. Moreover, the existence of cross-situational consistency demonstrates that raters were assessing stable constructs rather than state-relevant manifestations. However, it is not clear whether these composite ratings demonstrate divergent validity with one another and convergent validity with similar measures. To examine convergent validity, we examined zero-order and partial correlations (controlling for age, sex, and income-to-needs) between parental reports of their child's temperament on subscales of the CBQ administered as part of the larger longitudinal study and thin slice personality dimension ratings (Table 6). In both zero-order and partial correlations, thin slice ratings of extraversion were positively correlated with parental reports of approach and high intensity pleasure and negatively correlated with parental reports of shyness. Thin slice ratings of openness were also positively correlated with parental reports of approach and high intensity pleasure. In zero-order correlations, conscientiousness was negatively correlated with anger/frustration, and positively correlated with low intensity pleasure and smiling/laughter. Agreeableness was positively correlated with approach and low intensity pleasure. In partial correlations, thin slice ratings of neuroticism were negatively correlated with high intensity pleasure and positively correlated with shyness.

To examine divergent validity, we examined the correlations among the thin slice personality ratings (reported in Table 3). Extraversion was highly, positively correlated with openness but negatively correlated with each of the other personality dimensions. Agreeableness and conscientiousness were highly, positively correlated. Neuroticism was negatively correlated with each of the other personality dimensions. Given the high correlations between: (1) extraversion and openness and (2) agreeableness and conscientiousness, divergent validity was low for those dimensions.

DISCUSSION

The current study demonstrated support for the detection of reliable personality dimensions in the preschool period using observed thin slice ratings. Thin slice ratings of preschoolers showed within-task/inter-rater agreement, internal consistency, and cross-rater/cross-task agreement that was similar to that seen in past work sampling different developmental stages. These findings add to the burgeoning literature suggesting that such dimensions can be reliably observed during the preschool period. Thin slice ratings were highly consistent across tasks and slightly more consistent than what has been previously reported in adults. Further, personality dimensions in preschoolers were shown to be reliably observable by blind, unacquainted raters (e.g., not parents or teachers). These results provide additional

support for the use of the thin slice coding technique as a cost-effective, efficient way to assess personality dimensions using objective methods that are not biased by the parent's own personality or psychopathology.

We found evidence for some early emerging gender differences in ratings of personality dimensions. Girls were rated as higher on conscientiousness and agreeableness than boys, whereas boys were rated as higher on neuroticism. Similar gender differences during early and middle childhood on agreeableness and conscientiousness have been found in past work using parental reports (Soto, 2016; Van den Akker, Deković, Asscher, & Prinzie, 2014). Further, in adult studies, women tend to score higher on agreeableness than men (Weisberg, DeYoung, & Hirsh, 2011). Evidence of gender differences during the preschool period for agreeableness and conscientiousness suggests that these gender differences may emerge earlier than previously realized. This may reflect widespread social-cultural norms and influences that broadly shape personality development from a young age, or they may be early manifestations of biological sex differences in personality. In this study, girls were also rated as lower on neuroticism than boys. Substantial gender differences in neuroticism tend to emerge by late adolescence and persist into adulthood, with girls becoming increasingly prone to displaying more neuroticism than boys (Soto, 2016; Soto & Tackett, 2015), however gender differences in neuroticism have not been shown in children as young as the current sample. Thus, the findings describe here extend these prior results by demonstrating that during early childhood, boys may actually display more neuroticism than girls, particularly boys with clinically elevated levels of psychopathology who may also be more likely to display anger and irritability during the tasks. Over time and across maturation, girls increase in mean levels of neuroticism much like what is seen in symptoms of depression, which tend to be equivalent during childhood for boys and girls but rapidly increase for girls during adolescence.

These results also indicate evidence for agreement across the thin slice raters. Two unacquainted observers rating the same preschooler in different situations showed convergence (average ICC= .33). This is similar to the only other published study using a thin slice technique to rate older children's personality dimensions which showed that two unacquainted observers rating personality dimensions of older children had an average ICC of 0.43 (Tackett et al., 2016). We also examined the internal consistency across all raters (e.g., the equivalent of Chronbach's alpha for *k* raters assessing the same task) and found that the reliability of averaged ratings was fair across all dimensions (average ICC = .57), comparable with what has been reported in past research in older child samples. For instance, Tackett and colleagues (Tackett et al., 2016) reported the average ICC for the internal consistency across all raters was 0.69. Furthermore, agreement of different observers rating different tasks evidenced good consensus (average $r = .57$). In comparison, Tackett and colleagues (Tackett et al., 2016) reported an average $r = .43$. This suggests that independent thin slice ratings show the type of reliability that would be expected if ratings were tapping into valid underlying personality traits or dimensions. Although the validity of this approach has yet to be fully established.

Estimates of cross-task or situation consistency across the eight thin slice tasks were high, ranging from 0.76 to 0.88 with a mean of 0.82. Our cross-task consistency estimates were

highest for extraversion and lowest for neuroticism. These findings are markedly higher than past work using adult (Borkenau et al., 2004) and even older child samples (Tackett et al., 2016), despite a smaller number of thin slice tasks. However, the other work that has used the thin slice coding technique has also used significantly more raters than we used in our study (8-18 in our study vs. 45 and 120 raters in prior work). This may partially explain why our cross-task/situation consistency estimates were higher. Alternatively, there may potential developmental differences in cross-situational consistency in that preschoolers appear to be more consistent than older children and adults in the personality traits that they exhibit across different situations/tasks. For instance, estimates in Borkenau et al.'s (2004) study ranged from 0.29 to 0.61 with a mean of 0.43. In Tackett et al.'s study (2016), estimates ranged from 0.41 to 0.71 with a mean of 0.63. Although the long-term stability of personality traits is likely lower in childhood than in adulthood (Soto & Tackett, 2015), these findings indicate that within a single assessment, preschoolers may display higher levels of consistency in their behavior than older children or adults. As suggested by Tackett and colleagues (2017), preschoolers may be more likely to authentically display emotion, letting their "real" personality shine through. The preschool period represents a unique time of development where children are just beginning to understand the value of managing emotional expressions and using emotional display rules in social situations (Murray & Rosanbalm, 2017; Rosanbalm & Murray, 2017). Thus, during these interactions, preschoolers may be less likely than older children or adults to engage in the types of emotion and/or impression management that would influence a rater's evaluation of personality traits. For example, in adolescents, higher agreement was found in self-other ratings of visible traits, such as extraversion when compared to less visible traits like neuroticism. Further the authors found that there were fewer inter-judge differences in the development of more visible traits during adolescence than for less visible traits. Exploring and quantifying these potential developmental shifts in cross-situational consistency represents an exciting area for future research.

Although cross-situation consistency was relatively high, certain tasks appeared to be more useful for rating specific personality dimensions, and certain tasks had higher ICCs overall. Across all personality dimensions, the *Storytelling* task had the highest average rater/ single task ICCs, whereas the *Empty Box* task had the lowest average rater/ single task ICCs. The Storytelling task also had the highest ICCs for extraversion, neuroticism, and openness, whereas the Snack Delay task had the highest ICC's for agreeableness and conscientiousness. It is likely that the demands of each task may have contributed to the consistency and specific dimension agreement. In the Storytelling task, children are asked to tell a story about what he/she did the previous day while two experimenters listen to the child's description. While originally designed to assess for social inhibition and shyness, this task offers thin slice coders the ability to hear the child describe personal aspects of their life, in their own words, without the influence of an experimenter or object to interact with. Given the relatively unstructured nature of the task compared with other tasks in the assessment battery, more accurate reflections of the child's true personality may shine through. On the other hand, the Empty Box task asks children to unwrap a gift in front of an experimenter, but it soon becomes apparent to the child that the box is empty. Originally designed to elicit aspects of negative emotionality, this task had the lowest ICCs in our

battery. Perhaps children were able to enact social display rules to hide or control their true reaction, particularly in front of an unknown adult authority figure. Coders may then have had difficulty accurately rating personality dimensions. Alternatively, there may have been a more restricted range of child affect and behaviors during this task that led to ratings based on limited information. In the Snack Delay task, children are required to wait for an experimenter to ring a bell prior to eating a snack. During the waiting period, children used many strategies to distract themselves from the snack. These strategies may have offered our raters specific clues about the child's personality, leading to enhanced agreement. Given that this study aimed to provide evidence of initial reliably and validity of this approach in young children and to be most conservative, we chose to include the ratings from all of the tasks in our analyses, as we had no a priori hypotheses that one task would be "better" than another at eliciting reliable ratings. However, it is likely that excluding certain tasks, such as the Empty Box that appeared to be less consistently rated across coders would have increased our reliability estimates. This type of task selection represents an exciting potential for future research to determine the fewest number of tasks required to create reliable, consistent ratings as well as which tasks may be the most useful for rating specific personality dimensions. To answer these intriguing questions, future thin-slice research would benefit from continued research using existing LABTAB tasks as well as employing less structured tasks, such as free play.

Evidence for convergent and divergent validity was also found in thin slice ratings of preschoolers' personality dimensions. Thin slice ratings were positively correlated with parental reports of preschoolers' temperament. More specifically, thin slice ratings of extraversion were correlated with parental reports of high intensity pleasure and negatively correlated with parental reports of shyness. Thin slice ratings of neuroticism showed the opposite pattern and were negatively correlated with high intensity pleasure and positively correlated with shyness. Divergent validity tests suggest that raters were not easily able to differentiate between extraversion and openness, and between conscientiousness and agreeableness. These associations are not entirely surprising given that the adult personality literature finds support for a three factor models that combine the same traits together (Markon et al., 2005). Further, there is mixed evidence for the number of dimensions that emerge from parent-reported assessments of their child's personality (Measelle et al., 2005; Tackett et al., 2012). In contrast to suggestions that childhood personality assessed by expanding the Big Five into the Little Six (Soto, 2016), these findings may suggest that the factor structure is simpler in early childhood with high correlations often found between extraversion and openness, and between conscientiousness and agreeableness using parent report instruments. Alternatively, the tasks in the current study may not be well suited to differentiating between these traits. Future research may need to design observational tasks that provide better divergent validity.

These findings represent a novel contribution to the literature on preschool personality, as we are unaware of any empirical report that has investigated personality dimensions using this approach during this developmental period. Our results indicate that overall, preschoolers' personality dimensions were rated similarly across each task and by different observers. Many believe that early temperament dimensions gradually develop into the more dispositional personality traits not observable until later in development (McAdams &

Olson, 2010). However, our findings suggest that even during the preschool period, there appears to be some consistency and reliability in observable personality dimensions. Further, these ratings were completed by observers who were blind to both diagnostic status of each child, as well as his/her previous behavior from other tasks, further enhancing the integrity of the thin slice approach.

Limitations of the current study should be noted. First, 197 (65%) of preschoolers had psychiatric disorders at baseline and the sample was enriched specifically for preschool depressive disorders. Therefore, it is important to replicate these findings in a community-based sample to ensure that the findings do not represent a manifestation of a concurrent mental disorder rather than a stable personality feature. Second, neuroticism demonstrated lower variability across children (overall low means) and somewhat lower internal consistency, and as a result, our neuroticism findings should be interpreted with caution. Of note, neuroticism has been suggested to be the most difficult Big 5 dimension to code using a thin slice approach (Widiger & Costa Jr, 1994), especially in children (Borkenau et al., 2004). As such, the observational tasks may not have elicited high enough levels of neuroticism to be reliably observed. In addition, given that the sample was enriched for depression and the association between depression and neuroticism, there may have been less variance in neuroticism among participants to meaningfully rate. An additional limitation is that in each of the laboratory tasks used to code personality dimensions, an adult research assistant was present in the room. Perhaps some children behaved differently due to the presence of this adult. Future research would benefit from observing children using less formal structured tasks, such as free play to enhance the validity of these findings. Finally, our ratings of personality dimensions were conducted using single-items, which may not have the same level of nuance as other more detailed ratings, such as the Hierarchical Personality Inventory for Children (Mervielde & De Fruyt, 1999; Vollrath et al., 2016). Future work using the thin slice technique with young children should consider the benefits of more detailed ratings scales for observers.

Our evidence for validity was limited given the lack of direct overlap between the thin slice personality dimensions and temperament assessed via the CBQ. The effect sizes seen in these relations were small, and evidence of validity using a more direct comparison of parent or clinician assessments of personality is needed. Furthermore, much work needs to be done in terms of establishing validity for the thin slice approach, regardless of child age. Along the same lines, raters were not able to validly discern between each of the five personality dimensions. While this is quite common in other, standard personality assessments in young children (Tackett et al., 2012), this limitation may prevent researchers aiming to use the thin slice technique from collecting enough information about each of the five traits using this approach. We recommend using both the thin slice technique and a standard assessment of personality traits in order to capture the depth and subtleties of each dimension. Moving forward, additional work that takes a specific task by personality dimension approach in terms of coding will likely improve reliability and validity estimates. We suggest using a task or tasks that directly capture the dimensions of interest, such as a snack delay or similar task for conscientiousness, a positive affect eliciting task for extraversion and agreeableness, and/or a story telling/narrative task for most personality dimensions.

CONCLUSIONS

The current study used a large sample of clinically heterogeneous preschoolers with observational data across a series of laboratory tasks in order to determine whether Big Five personality dimensions could be reliably identified. Findings illustrate that personality dimensions can be reliably observed using the thin slice technique in children as young as preschool. Thin slice ratings demonstrated encouraging psychometric properties including within-task, inter-rater agreement and cross-situational consistency across the preschool developmental period. Further research is needed to replicate and extend this work using additional samples of preschoolers to investigate whether these personality dimensions in preschoolers are stable longitudinally, as well as continue to establish the validity of the thin slice approach.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors wish to thank the children and caregivers of the Preschool Depression Study for their time and dedication to this project. The authors also wish to thank the staff and students of the Early Emotional Development Program for their tireless efforts in the coding of preschool personality dimensions.

FUNDING

All phases of this study were supported by an NIH grant, R01 MH064769-06A1 (PI:s Barch and Luby). Dr. Gilbert's work was supported by K23 MH115074-01 (PI: Gilbert). Dr. Whalen's work was supported by NIH grants: K23 MH118426-01 (PI: Whalen) and L30 MH108015 (PI: Whalen).

REFERENCES

Abe JA (2005). The predictive validity of the Five-Factor Model of personality with preschool age children: A nine-year follow-up study. Journal of Research in Personality, 39(4), 423–442. 10.1016/j.jrp.2004.05.002

Abe JA, & Izard CE (1999). A longitudinal study of emotion expression and personality relations in early development. Journal of Personality and Social Psychology, 77(3), 566–577. [PubMed: 10510509]

Ambady N, & Rosenthal R (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. Psychological Bulletin, 111(2), 256–274.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). Washington, DC: American Psychiatric Association.

Asendorpf JB, & Denissen JJA (2006). Predictive validity of personality types versus personality dimensions from early childhood to adulthood: Implications for the distinction between core and surface traits. Merrill-Palmer Quarterly, 52(3), 486–513.

Borkenau P, Mauer N, Riemann R, Spinath FM, & Angleitner A (2004). Thin slices of behavior as cues of personality and intelligence. Journal of Personality and Social Psychology, 86(4), 599–614. 10.1037/0022-3514.86.4.599 [PubMed: 15053708]

Borkenau P, & Zaltauskas K (2009). Effects of self-enhancement on agreement on personality profiles. European Journal of Personality, 23(2), 107–123. 10.1002/per.707

Carcone AI, Naar S, Eggly S, Foster T, Albrecht TL, & Brogan KE (2015). Comparing thin slices of verbal communication behavior of varying number and duration. Patient Education and Counseling, 98(2), 150–155. 10.1016/j.pec.2014.09.008 [PubMed: 25441095]

Carney DR, Colvin CR, & Hall JA (2007). A thin slice perspective on the accuracy of first impressions. Journal of Research in Personality, 41(5), 1054–1072. 10.1016/j.jrp.2007.01.004

Caspi A, Roberts BW, & Shiner RL (2005). Personality development: Stability and change. Annual Review of Psychology, 56(1), 453–484. 10.1146/annurev.psych.55.090902.141913

Connelly BS, & Ones DS (2010). Another perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. Psychological Bulletin, 136(6), 1092. [PubMed: 21038940]

De Pauw SSW (2017). Childhood personality and temperament The Oxford Handbook of the Five Factor Model. New York: Oxford University Press, pages 243–280.

De Pauw SSW, & Mervielde I (2010). Temperament, personality and developmental psychopathology: A review based on the conceptual dimensions underlying childhood traits. Child Psychiatry & Human Development, 41(3), 313–329. 10.1007/s10578-009-0171-8 [PubMed: 20238477]

De Pauw SSW, Mervielde I, & Leeuwen KGV (2009). How are traits related to problem behavior in preschoolers? Similarities and contrasts between temperament and personality. Journal of Abnormal Child Psychology, 37(3), 309–325. 10.1007/s10802-008-9290-0 [PubMed: 19165590]

Durbin EC, & Wilson S (2012). Convergent validity of and bias in maternal reports of child emotion. Psychological Assessment, 24(3), 647–660. 10.1037/a0026607 [PubMed: 22149326]

Dyson MW, Olino TM, Durbin CE, Goldsmith HH, Bufferd SJ, Miller AR, & Klein DN (2015). The structural and rank-order stability of temperament in young children based on a laboratory-observational measure. Psychological Assessment, 27(4), 1388–1401. 10.1037/pas0000104 [PubMed: 25894709]

Dyson MW, Olino TM, Durbin CE, Goldsmith HH, & Klein DN (2012). The structure of temperament in preschoolers: A two-stage factor analytic approach. Emotion, 12(1), 44–57. 10.1037/a0025023 [PubMed: 21859196]

Goldberg LR (1993). The structure of phenotypic personality traits. The American Psychologist, 48(1), 26–34. [PubMed: 8427480]

Goldsmith H, Reilly J, Lemery K, Longley S, & Prescott A (1999). Laboratory Temperament Assessment Battery: Preschool Version (version 1.0, Technical Report). Department of Psychology, University of Wisconsin, Madison, WI.

Goldstein BL, Shankman SA, Kujawa A, Torpey-Newman DC, Dyson MW, Olino TM, & Klein DN (2019). Positive and negative emotionality at age 3 predicts change in frontal eeg asymmetry across early childhood. Journal of Abnormal Child Psychology, 47(2), 209–219. 10.1007/s10802-018-0433-7 [PubMed: 29687430]

Grist CL, & McCord DM (2010). Individual differences in preschool children: Temperament or personality? Infant and Child Development, 19(3), 264–274. 10.1002/icd.663

Halverson CF, Havill VL, Deal J, Baker SR, Victor JB, Pavlopoulos V, … Wen L (2003). Personality structure as derived from parental ratings of free descriptions of children: The inventory of child individual differences. Journal of Personality, 71(6), 995–1026. [PubMed: 14633056]

Herzhoff K, Kushner SC, & Tackett JL (2017). Personality development in childhood In Specht J (Ed.), Personality Development Across the Lifespan (pp. 9–23). Academic Press 10.1016/B978-0-12-804674-6.00002-8

Hirschmann N, Kastner-Koller U, Deimann P, Schmelzer M, & Pietschnig J (2018). Reliable and valid coding of thin slices of video footage: Applicability to the assessment of mother-child interactions. Journal of Psychopathology and Behavioral Assessment, 40(2), 249–258. 10.1007/s10862-017-9630-x [PubMed: 29937622]

Jackson JJ, Walton KE, Harms PD, Bogg T, Wood D, Lodi-Smith J, … Roberts BW (2009). Not all conscientiousness scales change alike: A multimethod, multisample study of age differences in the facets of conscientiousness. Journal of Personality and Social Psychology, 96(2), 446–459. 10.1037/a0014156 [PubMed: 19159142]

Koo TK, & Li MY (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155–163. 10.1016/j.jcm.2016.02.012 [PubMed: 27330520]

Kopala-Sibley DC, Olino T, Durbin E, Dyson MW, & Klein DN (2018). The stability of temperament from early childhood to early adolescence: A multi-method, multi-informant examination. European Journal of Personality, 32(2), 128–145. 10.1002/per.2151 [PubMed: 30858648]

Kraus MW, & Keltner D (2009). Signs of socioeconomic status: A thin-slicing approach. Psychological Science, 20(1), 99–106. 10.1111/j.1467-9280.2008.02251.x [PubMed: 19076316]

Kryski KR, Olino TM, Dyson MW, Durbin CE, Klein DN, & Hayden EP (2018). Associations between observed temperament in preschoolers and parent psychopathology. Personality and Mental Health, 12(2), 131–144. 10.1002/pmh.1406 [PubMed: 29160041]

Lamb ME, Chuang SS, Wessels H, Broberg AG, & Hwang CP (2002). Emergence and construct validation of the big five factors in early childhood: A longitudinal analysis of their ontogeny in Sweden. Child Development, 73(5), 1517–1524. [PubMed: 12361316]

Luby JL, Belden AC, Pautsch J, Si X, & Spitznagel E (2009). The clinical significance of preschool depression: Impairment in functioning and clinical markers of the disorder. Journal of Affective Disorders, 112(1–3), 111–119. 10.1016/j.jad.2008.03.026 [PubMed: 18486234]

Luby JL, Heffelfinger A, Koenig-McNaught AL, Brown K, & Spitznagel E (2004). The Preschool Feelings Checklist: A brief and sensitive screening measure for depression in young children. Journal of the American Academy of Child & Adolescent Psychiatry, 43(6), 708–717. 10.1097/01.chi.0000121066.29744.08 [PubMed: 15167087]

Luby JL, Si X, Belden AC, Tandon M, & Spitznagel E (2009). Preschool depression: Homotypic continuity and course over 24 months. Archives of General Psychiatry, 66(8), 897–905. 10.1001/archgenpsychiatry.2009.97 [PubMed: 19652129]

Markon KE, Krueger RF, & Watson D (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. Journal of Personality and Social Psychology, 88(1), 139–157. 10.1037/0022-3514.88.1.139 [PubMed: 15631580]

McAdams DP, & Olson BD (2010). Personality development: Continuity and change over the life course. Annual Review of Psychology, 61(1), 517–542. 10.1146/annurev.psych.093008.100507

McCrae RR, Costa PT, Ostendorf F, Angleitner A, Hrebícková M, Avia MD, … Smith PB (2000). Nature over nurture: Temperament, personality, and life span development. Journal of Personality and Social Psychology, 78(1), 173–186. [PubMed: 10653513]

McLoyd VC (1998). Socioeconomic disadvantage and child development. American Psychologist, 53(2), 185–204. 10.1037/0003-066X.53.2.185

Measelle JR, John OP, Ablow JC, Cowan PA, & Cowan CP (2005). Can children provide coherent, stable, and valid self-reports on the big five dimensions? A longitudinal study from ages 5 to 7. Journal of Personality and Social Psychology, 89(1), 90–106. 10.1037/0022-3514.89.1.90 [PubMed: 16060748]

Mervielde I, & De Fruyt F (1999). Construction of the Hierarchical Personality Inventory for Children (HiPIC) In Deary IJ, De Fruyt F, & Ostendorf F (Eds). Personality psychology in Europe (Vol. 7, pp. 107–127). Tilburg, Netherlands: Tilburg University Press

Murphy NA, Hall JA, Ruben MA, Frauendorfer D, Schmid Mast M, Johnson KE, & Nguyen L (2019). Predictive validity of thin-slice nonverbal behavior from social interactions. Personality and Social Psychology Bulletin, 45(7), 983–993. 10.1177/0146167218802834 [PubMed: 30400748]

Murphy NA, Hall JA, Schmid Mast M, Ruben MA, Frauendorfer D, Blanch-Hartigan D, … Nguyen L (2015). Reliability and validity of nonverbal thin slices in social interactions. Personality and Social Psychology Bulletin, 41(2), 199–213. 10.1177/0146167214559902 [PubMed: 25488925]

Murray DW, & Rosanbalm K (2017). Promoting Self-Regulation in Adolescents and Young Adults: A Practice Brief. OPRE Report 2015-82. Retrieved from https://eric.ed.gov/?id=ED594226

Olino TM, Klein DN, Dyson MW, Rose SA, & Durbin CE (2010). Temperamental emotionality in preschool-aged children and depressive disorders in parents: Associations in a large community sample. Journal of Abnormal Psychology, 119(3), 468–478. 10.1037/a0020112 [PubMed: 20677836]

Prime H, Perlman M, Tackett JL, & Jenkins JM (2014). Cognitive sensitivity in sibling interactions: Development of the construct and comparison of two coding methodologies. Early Education and Development, 25(2), 240–258. 10.1080/10409289.2013.821313

Rosanbalm KD, & Murray DW (2017). Promoting Self-Regulation in the First Five Years: A Practice Brief. OPRE Brief 2017-79. Retrieved from https://eric.ed.gov/?id=ED583624

Rothbart MK, Ahadi SA, Hershey KL, & Fisher P (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. Child Development, 72(5), 1394–1408. 10.1111/1467-8624.00355 [PubMed: 11699677]

Shiner RL, Masten AS, & Roberts JM (2003). Childhood personality foreshadows adult personality and life outcomes two decades later. Journal of Personality, 71(6), 1145–1170. 10.1111/1467-6494.7106010 [PubMed: 14633061]

Slepian ML, Bogart KR, & Ambady N (2014). Thin-slice judgments in the clinical context. Annual Review of Clinical Psychology, 10(1), 131–153. 10.1146/annurev-clinpsy-090413-123522

Soto CJ (2016). The little six personality dimensions from early childhood to early adulthood: Mean-level age and gender differences in parents' reports. Journal of Personality, 84(4), 409–422. 10.1111/jopy.12168 [PubMed: 25728032]

Soto CJ, & Tackett JL (2015). Personality traits in childhood and adolescence: Structure, development, and outcomes. Current Directions in Psychological Science, 24(5), 358–362. 10.1177/0963721415589345

Tackett JL, Herzhoff K, Kushner SC, & Rule N (2016). Thin slices of child personality: Perceptual, situational, and behavioral contributions. Journal of Personality and Social Psychology, 110(1), 150–166. 10.1037/pspp0000044 [PubMed: 25938703]

Tackett JL, Lang JWB, Markon KE, & Herzhoff K (2019). A correlated traits, correlated methods model for thin-slice child personality assessment. Psychological Assessment, 31(4), 545–556. 10.1037/pas0000635 [PubMed: 30869957]

Tackett JL, Slobodskaya HR, Mar RA, Deal J, Halverson CF, Baker SR, … Besevegis E (2012). The hierarchical structure of childhood personality in five countries: Continuity from early childhood to early adolescence. Journal of Personality, 80(4), 847–879. 10.1111/j.1467-6494.2011.00748.x [PubMed: 22091829]

Tackett JL, Smack AJ, Herzhoff K, Reardon KW, Daoud S, & Granic I (2017). Measuring child personality when child personality was not measured: Application of a thin-slice approach. Personality and Mental Health, 11(1), 4–13. 10.1002/pmh.1351 [PubMed: 27670747]

Van den Akker AL, Deković M, Asscher J, & Prinzie P (2014). Mean-level personality development across childhood and adolescence: A temporary defiance of the maturity principle and bidirectional associations with parenting. Journal of Personality and Social Psychology, 107(4), 736–750. 10.1037/a0037248 [PubMed: 25133720]

Vazire S, & Carlson EN (2011). Others sometimes know us better than we know ourselves. Current Directions in Psychological Science, 20(2), 104–108. 10.1177/0963721411402478

Victor JB, Rothbart MK, Baker SR, & Tackett JL (2016). Temperamental components of the developing personality In The Wiley Handbook of Personality Assessment (pp. 44–58). 10.1002/9781119173489.ch4

Vollrath ME, Hampson SE, & Torgersen S (2016). Constructing a short form of the hierarchical personality inventory for children (HiPIC): The HiPIC-30. Personality and Mental Health, 10(2), 152–165. 10.1002/pmh.1334 [PubMed: 27120426]

Weisberg YJ, DeYoung CG, & Hirsh JB (2011). Gender differences in personality across the ten aspects of the big five. Frontiers in Psychology, 2 10.3389/fpsyg.2011.00178

Widiger TA, & Costa PT Jr (1994). Personality and personality disorders. Journal of Abnormal Psychology, 103(1), 78. [PubMed: 8040485]

Wilson S, Schalet BD, Hicks BM, & Zucker RA (2013). Identifying early childhood personality dimensions using the california child q-set and prospective associations with behavioral and psychosocial development. Journal of Research in Personality, 47(4). 10.1016/j.jrp.2013.02.010

Zupančič M, Podlesek A, & Kavčič T (2006). Personality types as derived from parental reports on 3-year-old. European Journal of Personality, 20(4), 285–303. 10.1002/per.584

Table 1.

Thin Slice Laboratory Temperament Assessment Battery (LabTAB) Tasks

Task Name	Description	Temperament Evoked
1. Box Empty	Child given a wrapped gift to unwrap, however the wrapped box was empty	Negative emotionality
2. Impossibly Perfect Circles	Child instructed to draw circles on blank paper, each circle was critiqued and the child asked to draw another circle	Negative emotionality; Persistence
3. Popping Bubbles	Child and experimenter play together with a bubble-making toy	Positive affect; Interest
4. Picture Tearing	Child shown favorite picture of experimenter and then instructed to tear picture by another experimenter	Compliance; Guilt
5. Snack Delay	Child instructed to wait for experimenter to ring bell prior to eating a snack and experimenter delayed	Inhibitory control
6. Storytelling	Child instructed to tell a story about what he/she did yesterday standing in front of two experimenters	Social inhibition; Shyness
7. Transparent Box	Child given a desirable toy inside a locked transparent box and given incorrect ring of keys to open the box	Frustration; Persistence; Interest
8. Tea Cups	Child and experimenter have a tea party and child is given a teacup with a faulty handle	Guilt; Negative emotionality

Table 2.

Descriptions of Personality Dimensions Provided to Coders

Personality Dimension	Description
Extraversion	talkative, assertive, active, excitement-seeking, and fun-loving
Agreeable	trusting, straight-forward, helpful, easy-going, and modest
Conscientious	deliberate, orderly, competent, dutiful, and achievement-striving
Emotional	anxious, depressed, self-conscious, impulsive and vulnerable
Curious	non-conforming, seeks novelty and fantasy, and open to new ideas and values

Likert Rating

1- not at all, 2- a little, 3-moderately, 4- quite a bit, 5-extremely

Table 3.

Means, Standard Deviations, and Intercorrelations between Measures

	1.	2.	3.	4.	5.	6.	7.	8.
1. Child Age	5.36 (0.86)							
2. Gender	−0.05	155 male						
3. Income to Needs	0.04	−0.00	2.10 (1.17)					
4. Extraversion	−0.10 $^{+}$	0.00	−0.10 $^{+}$	2.96 (0.54)				
5. Agreeableness	0.25 ***	0.19 **	0.22 ***	−0.32 **	3.38 (0.39)			
6. Conscientiousness	0.33 ***	0.21 ***	0.37 ***	−0.34 **	0.81 **	3.09(0.43)		
7. Neuroticism	−0.09	−0.21 ***	0.06	−0.15 *	−0.42 **	−0.24 **	2.13(0.29)	
8. Openness	−0.19 **	−0.05	−0.01	0.78 **	−0.21 **	−0.30 **	−0.17 **	2.79 (0.35)

Note. All personality dimension ratings were made on a 5-point scale. N = 299

M(SD)

$^{+}$p<.10;

* $p < .05$;

** $p < .01$;

*** $p < .001$

Table 4.

Intraclass Correlation Coefficients Among Raters and Tasks of Thin-Slice Videotaped Personality Dimensions

Intraclass correlation coefficients (ICCs)				
Personality Dimension	**Single rater/ single task (1,1)**	**Average rater/ single task (1,3)**	**Cross rater/ cross task (2,1)**	**Cross-situational consistency**
Extraversion	.53	.77	.67	.87
Agreeableness	.30	.56	.49	.83
Conscientiousness	.34	.60	.47	.85
Neuroticism	.25	.48	.37	.76
Openness	.24	.46	.38	.82

Table 5.

Intraclass Correlation Coefficients (1,1) Among Raters (average rater/ single task) for each Task of Thin-Slice Videotaped Personality Dimensions

	Task							
Personality Dimension	**Empty Box**	**Impossible Circles**	**Popping Bubbles**	**Picture Tearing**	**Snack Delay**	**Storytelling**	**Transparent Box**	**Teacups**
Extraversion	.41	.46	.41	.63	.58	.65	.56	.56
Agreeableness	.31	.32	.31	.19	.42	.26	.30	.28
Conscientiousness	.12	.44	.36	.35	.53	.34	.30	.28
Neuroticism	.28	.31	.14	.14	.13	.37	.32	.32
Openness	.04	.28	.14	.22	.26	.39	.29	.28

Table 6.

Zero-order and Partial Correlations between Parental Ratings of Temperament and Thin Slice Ratings of Personality Dimensions

	Thin Slice Ratings of Personality Dimensions				
Zero-Order Correlation	**Extraversion**	**Openness**	**Conscientiousness**	**Neuroticism**	**Agreeableness**
Anger/Frustration	0.04	0.01	−0.13*	0.05	−0.11
Approach	0.17**	0.14**	0.02	−0.05	0.32**
Falling Reactivity/Soothability	−0.00	0.00	0.08	−0.03	0.07
Fear	0.01	0.01	0.07	−0.05	0.06
High Intensity Pleasure	0.21**	0.20**	−0.10	−0.08	−0.04
Low Intensity Pleasure	0.09	0.08	0.14*	−0.14*	0.15*
Sadness	0.12	0.11	0.07	−0.04	0.02
Shyness	−0.18**	−0.12	0.01	0.10	0.01
Smiling and Laughter	0.10	0.07	0.13*	0.03	0.10
Partial Correlation	**Extraversion**	**Openness**	**Conscientiousness**	**Neuroticism**	**Agreeableness**
Anger/Frustration	0.02	0.03	−0.06	0.02	−0.04
Approach	0.20**	0.17*	−0.02	−0.08	0.01
Falling Reactivity/Soothability	−0.00	−0.02	0.11	0.01	0.08
Fear	0.01	0.03	0.06	−0.04	0.05
High Intensity Pleasure	0.22**	0.21**	−0.03	−0.17*	0.03
Low Intensity Pleasure	0.11	0.09	0.11	−0.09	0.13
Sadness	0.12	0.12	0.02	−0.00	−0.03
Shyness	−0.19**	−0.12	−0.01	0.14*	−0.01
Smiling and Laughter	0.12	0.07	0.16*	0.03	0.10

Partial correlations account for child age, sex, and family income-to-needs ratio

*Correlation is significant at the 0.05 level (2-tailed);

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