

The M5–PS–35: A Five-Factor Personality Questionnaire for Preschool Children

CATHY L. GRIST,¹ ALAN SOCHA,² AND DAVID M. MCCORD³

¹*Department of Human Services, Western Carolina University*

²*Office of Institutional Planning and Effectiveness, Western Carolina University*

³*Department of Psychology, Western Carolina University*

The Five-factor theory of personality (FFT) has pervaded personality research in recent years. Although many reliable and valid measurement instruments exist for use with adults, adolescents, and even elementary-age children, there is a lack of available 5-factor measurement tools for use with preschool children. This article expands on previous work developing the M5–PS, a rating form for preschool children designed to be completed by classroom teachers or caregivers. A total of 621 children were rated by their teachers on the 90-item working form of the M5–PS. Through a combination of empirical and rational scale refinement methods, the number of items has been reduced to 35, yielding a revised instrument, the M5–PS–35, with substantially improved construct validity and scale internal consistency. Potential changes in external validity were evaluated by comparative reanalysis of an existing data set.

The Five-factor theory of personality (FFT) has clearly become the dominant paradigm for studying personality in adults (Gosling & John, 1999; Marsh et al., 2010), and to an increasing degree this framework is being applied to children, even very young ones (Deal, Halverson, Martin, Victor, & Baker, 2007; Kohnstamm, Halverson, Mervielde, & Havill, 1998). A major tool of researchers working with adult subjects is the self-report questionnaire, a tool not available to researchers working with young children. The most frequent substitute with child populations is the observer rating form, usually completed by parents or teachers. However, most such instruments are either grounded in temperament theories or are focused on specific pathological syndromes (e.g., Achenbach & Rescorla, 2000). This article describes the development and psychometric characteristics of an FFT questionnaire for preschool children, specifically the refinement of a 90-item working version of the instrument to a new 35-item version that exhibits improved reliability and construct validity.

Over the past 20 to 30 years the FFT has gradually achieved prominence in the field of personality psychology (Digman, 1990; Goldberg, 1993). Although not without critics (e.g., Block, 1995), the FFT has been shown as a useful perspective across a wide range of cultures, age groups, and even species (Costa & McCrae, 1992a; Gosling & John, 1999).

The downward extension of this theoretical framework to younger individuals has been somewhat slower, but recent years have shown much published research applying the FFT to adolescents and children (e.g., Costa & McCrae, 1994; Digman & Shmelyov, 1996; Kohnstamm et al., 1998). Evidence is rapidly accumulating that this theoretical framework is indeed very ap-

plicable to the preschool age range and could supplant traditional temperamental models in describing individual differences in young children (see, e.g., Abe & Izard, 1999; Asendorpf & Denissen, 2006; De Pauw, Mervielde, & Van Leeuwen, 2009; Grist & McCord, 2010; Halverson et al., 2003). There has also been research with self-report measures developed for young children ages 5 to 7 (Measelle, John, Ablow, Cowan, & Cowan, 2005). However, it is difficult to use self-report measures reliably with preschool children ages 3 and 4 due to the nature of language and cognitive development of this age group.

Some researchers caution against simply assuming that adult models of individual differences, such as the FFT, automatically apply to children as well. See, for example, Kohnstamm et al. (1998) for a discussion of these concerns. However, the five factors do have strong roots in the evolutionary psychology literature, where they are seen as evolved psychological mechanisms that are basic to human nature (e.g., D. M. Buss, 1991; Nettle, 2006). Consistent with this perspective, Gosling and John (1999) presented data showing the presence of all five factors in the closest relative of humans, the common chimpanzee. Seminal work by Digman, based on data collected from 1959 to 1967, reanalyzed by Goldberg (2001), also provides strong support for the basic idea that the FFT is applicable and useful in describing individual differences in children. Briefly, teacher ratings of more than 2,500 children were collected by Digman and colleagues on two Hawaiian islands. Digman was arguing for a model of child personality that included at least 10 different factors and thus included a wide range of variables. Note that the time period of this research predates the emergence of the FFT as a dominant paradigm. In data from all six samples, which varied in size from 100 to 885 children from first to sixth grade, the five-factor structure emerged clearly. In addition, Goldberg argued that there was no evidence of broad factors beyond these five. Further, he documented the assertion that Digman's child data contributed substantially to the clarification, emergence, and dominance of the FFT. Finally, although Kohnstamm et al. (1998) argued on a conceptual level that we should be hesitant to apply adult models to children, the empirical data from their

Received April 5, 2011; Revised August 16, 2011.

Alan Socha is now at the Center for Assessment and Research Studies, James Madison University.

Address correspondence to Cathy L. Grist, Department of Human Services, Western Carolina University, 301 Killian Building, Cullowhee, NC 28723; Email: clgrist@wcu.edu

large, methodologically excellent, cross-cultural study yields possibly the most compelling data to date that the FFT is the most appropriate framework for studying child personality. Researchers in seven different countries interviewed parents in their homes, asking the simple open-ended question, "Tell me about your child." Data were recorded, transcribed, and coded. Of the total of 24,000 descriptors provided by parents, approximately 85% fit more easily into the five factors of the FFT than into any other of the 14 categories available to coders. The original 14 categories included traditional temperament concepts as well as other variables suggested by the developmental psychology literature. That is, the FFT was not at all an a priori organizational framework, but rather one that emerged empirically from the rigorously analyzed qualitative data.

On a practical level, one constraining factor has been a relative dearth of rating forms based on the FFT and designed for preschool children. Personality researchers working with adult subjects have many well-validated instruments, including the Revised NEO Personality Inventory (NEO PI-R) and the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992b), the Big Five Inventory (John & Srivastava, 1999), the Hogan Personality Inventory (Hogan & Hogan, 1992), and public-domain proxy scales for all of these instruments, and more, available through Goldberg's (1999) International Personality Item Pool (IPIP) project Web site (see also International Personality Item Pool, 2001). Psychologists studying personality in preschool children typically find themselves with potentially useful theory but no measurement tool.

Historically, child development specialists have viewed preschool children through the lens of temperament constructs, which are generally seen as having a strong neurobiological and genetic basis (Mervielde, De Clercq, De Fruyt, & Van Leeuwen, 2005). Assessment instruments have thus been based primarily on temperament theories (e.g., A. H. Buss & Plomin, 1975; Rothbart & Derryberry, 1981; Thomas & Chess, 1977). In contrast, individual differences in adults are most commonly described as personality characteristics or traits, which might be rooted in temperament, but are also influenced by environment and experience. As noted earlier, over the last decade or so the use of personality theory, specifically the FFT, to assess adolescents and school-age children has increased significantly (De Clercq & De Fruyt, 2003; De Fruyt, Mervielde, Hoekstra, & Rolland, 2000; Eysenck, Makaremi, & Barrett, 1994). Furthermore, there has been a downward extension of the FFT to incorporate individual differences in preschool children as well. Although several FFT-based measures have been developed for use with children and adolescents, they are not suitable for preschoolers due to significant developmental differences between young children ages 3 to 5 and children 6 and older. Two notable exceptions are the Hierarchical Personality Inventory for Children (HPiC; Mervielde & De Fruyt, 2002) and the Inventory of Children's Individual Differences (ICID; Deal et al., 2007; Halverson et al., 2003).

The ICID original form consisted of 108 items, and the 2007 revision has 50 items. This instrument derives from the landmark research of Kohnstamm et al. (1998), described earlier, in which they obtained "free descriptions" from parents about their children from seven different countries. These narrative data were analyzed through rigorous qualitative research techniques and about 85% of the coded parent comments were reliably categorized in conceptual groupings based on the FFT.

The ICID exhibits good reliability and validity and is available to researchers from the authors.

In this article we describe the development of another questionnaire for preschoolers, the M5-PS-35, starting from a different place. Briefly, whereas Halverson and colleagues approached the task of systematizing the study of child personality qualitatively and relatively atheoretically, through parental free descriptions, we base our own efforts squarely in the historical lineage of the FFT.

Development of the M5-PS

We began by adopting the FFT as our conceptual framework for child personality. Thus, the first step in developing the M5-PS questionnaire was to identify a broad item pool that contained a breadth of content sufficient to span the five factors. Goldberg's (1999) IPIP project provides at this time 2,413 personality items of similar format, freely available on the IPIP Web site (International Personality Item Pool, 2001). Through correlational analysis, these items have been used to form "proxy" scales for many published personality instruments as well as unpublished research scales. Of these 2,413 items, 336 have been specifically correlated with domains and facets of Costa and McCrae's NEO PI-R. Three different proxy instruments are presented by Goldberg (1999). First, the 10 items correlating most highly with each of the six facets of each of the five domains were identified, producing a 300-item set that allows measurement of each of the 30 facets, with scores that can be summed across facets to measure the five broad domains (with 60 items each). A second proxy instrument was developed by identifying the 20 items with the highest correlations with each of the five domains, producing a 100-item questionnaire that yields the five domain scores only. Finally, a shorter form was developed using the 10 items with the highest correlations with each domain, resulting in a 50-item set. Of course, all items in the 50-item set also occur on the 100-item set. However, because the 300-item set was developed by targeting facets rather than domains, some items were included on the 50- and 100-item sets that were not included on the 300-item set. Thus, it takes 336 items to completely capture all of the identified high correlates with the NEO PI-R (see Goldberg et al., 2006, for a more detailed presentation of the development of the IPIP proxy scales).

These 336 items are, of course, adult oriented. They have been used satisfactorily with older and midadolescents, but not with children. Our next step, then, was to review each of the 336 items, eliminating many and rewriting many more, to create an item set that would be appropriate for preschool children. The IPIP items are written in first person with pronoun implied. We rewrote all items in implied third person, which in some cases required modifying pronouns to gender-neutral forms (e.g., himself/herself). Many items had content that was not relevant to this age group and were thus discarded, such as "Am politically liberal," "Use flattery to get ahead," "Am not interested in abstract ideas," and "Pay my bills on time." With regard to content, most rewritten items are very similar to the original IPIP items. This step resulted in 158 items that were seen as appropriate for the age range, written in a form to be rated by a parent or preschool teacher. This experimental form was labeled the M5-PS-X.

The next step was to have the 158 items rated by a panel of judges as to their relevance for the preschool age group. A

group consisting of 3 early childhood educators, 10 preschool teachers, and 10 parents of preschool-age children provided the ratings. The raters were asked to assess each item independently using a five-category Likert-type scale, labeled *Totally Irrelevant*, *Somewhat Irrelevant*, *Neither*, *Somewhat Relevant*, and *Very Relevant*. Raters were asked to think of a specific child or group of children while determining each item's relevance or appropriateness for the particular quality listed. Items were then sorted by average "relevance" rating using data from all raters, and 90 items were selected for the final questionnaire. The final number was chosen partially with regard to length of total questionnaire; a preschool teacher should be able to complete

a 90-item scale in 10 minutes. Adjustments were made to include 18 items from each of the five personality domains, based on original item association with the adult scale. Specifically, when ranked by judged relevance, we were able to identify 18 Extraversion items by the 74th item. We then skipped subsequent Extraversion items, selecting for other scales. We were able to identify 18 items for each scale, for a total of 90 items, by the time we reached the item ranked 102. The mean relevance rating was still on the "relevant" side of neutral for all items selected.

The 90-item scale, or M5-PS-90 (see Table 1), was then utilized in a variety of data-collection efforts and exploratory

TABLE 1.—M5-PS-90 items.

Item No.	F/R	Text	M5 Factor/Facet	Item No.	F/R	Text	M5 Factor/Facet
1	+	Worries about things	N1-Anxiety	46	–	Does not like the idea of change	O4-Adventurousness
2	+	Has a vivid imagination	O1-Imagination	47	+	Loves action	E5-Excitement-seeking
3	–	Distrusts people	A1-Trust	48	+	Feels comfortable around other people	E1-Friendliness
4	+	Completes tasks successfully	C1-Self-efficacy	49	+	Trusts what people say	A1-Trust
5	+	Gets angry easily	N2-Anger	50	+	Loves order and regularity	C2-Orderliness
6	+	Takes charge	E3-Assertiveness	51	+	Loves to help others	A3-Altruism
7	–	Seldom gets emotional	O3-Emotionality	52	–	Is a creature of habit	O4-Adventurousness
8	–	Breaks rules	C3-Dutifulness	53	–	Yells at people	A4-Cooperation
9	+	Is easily intimidated	N4-Self-consciousness	54	+	Plunges into tasks with all their heart	C4-Achievement Striving
10	+	Makes friends easily	E1-Friendliness	55	+	Has a rich vocabulary	O5-Intellect
11	+	Trusts others	A1-Trust	56	–	Knows the answers to many questions	A5-Modesty
12	+	Gets irritated easily	N2-Anger	57	–	Knows how to cope	N6-Vulnerability
13	+	Likes music	O2-Artistic Interests	58	+	Gets stressed out easily	N1-Anxiety
14	+	Experiences emotions intensely	O3-Emotionality	59	+	Acts comfortably with others	E1-Friendliness
15	+	Tries to follow the rules	C3-Dutifulness	60	+	Enjoys being part of a group	E2-Gregariousness
16	+	Is always busy	E4-Activity Level	61	–	Leaves his/her belongings around	C2-Orderliness
17	–	Prefers to stick with things that he/she knows	O4-Adventurousness	62	+	Tries to influence others	E3-Assertiveness
18	+	Is easy to satisfy	A4-Cooperation	63	+	Is concerned about others	A3-Altruism
19	+	Likes to solve complex problems	O5-Intellect	64	+	Tells the truth	C3-Dutifulness
20	+	Radiates joy	E6-Cheerfulness	65	+	Is interested in many things	O4-Adventurousness
21	–	Jumps into things without thinking	C6-Cautiousness	66	+	Involves others in what he/she is doing	E2-Gregariousness
22	+	Tries to excel at what they do	C1-Self-efficacy	67	+	Has frequent mood swings	N3-Depression
23	–	Is indifferent to the feelings of others	A3-Altruism	68	–	Experiences very few emotional highs and lows	O3-Emotionality
24	–	Is comfortable in unfamiliar situations	N4-Self-consciousness	69	–	Does the opposite of what is asked	C3-Dutifulness
25	+	Is always on the go	E4-Activity Level	70	–	Insults people	A4-Cooperation
26	–	Dislikes changes	O4-Adventurousness	71	–	Has difficulty starting tasks	C5-Self-discipline
27	+	Can't stand confrontations	A4-Cooperation	72	+	Loses his/her temper	N2-Anger
28	+	Has a lot of fun	E6-Cheerfulness	73	+	Likes to begin new things	O4-Adventurousness
29	+	Is afraid of many things	N1-Anxiety	74	–	Gets back at others	A4-Cooperation
30	+	Loves to daydream	O1-Imagination	75	+	Gets overwhelmed by emotions	N6-Vulnerability
31	–	Is wary of others	A1-Trust	76	+	Laughs aloud	E6-Cheerfulness
32	+	Sticks to the rules	A2-Morality	77	+	Suffers from others' sorrows	A6-Sympathy
33	–	Feels comfortable with him/herself	N3-Depression	78	–	Acts without thinking	C6-Cautiousness
34	+	Tries to lead others	E3-Assertiveness	79	–	Adapts easily to new situations	N1-Anxiety
35	–	Is not easily affected by his/her emotions	O3-Emotionality	80	–	Doesn't see the consequences of things	C1-Self-efficacy
36	–	Likes to take his/her time	E4-Activity Level	81	–	Is able to stand up for him/herself	N4-Self-consciousness
37	+	Works hard	C4-Achievement-striving	82	–	Makes him/herself the center of attention	A5-Modesty
38	+	Seeks adventure	E5-Excitement-seeking	83	+	Amuses his/her friends	E6-Cheerfulness
39	+	Becomes overwhelmed by events	N6-Vulnerability	84	+	Sympathizes with others' feelings	O (No facet association)
40	–	Is relaxed most of the time	N1-Anxiety	85	+	Is easily frustrated	N (No facet association)
41	–	Does not understand things	C1-Self-efficacy	86	+	Respects others	O (No facet association)
42	+	Gets upset easily	N2-Anger	87	–	Messes things up	A (No facet association)
43	–	Does not like crowded events	E2-Gregariousness	88	–	Is demanding	O (No facet association)
44	–	Knows how to get around the rules	A2-Morality	89	+	Starts conversations	C (No facet association)
45	+	Wants everything to be "just right"	C2-Orderliness	90	+	Finishes what he/she starts	C (No facet association)

Note. Coding direction (F/R): + means that the responses are coded in a positive direction; – means that they are reverse-coded.

projects. At this initial stage, each item was assumed to represent the FFT domain designated by its original IPIP association. Preliminary results from the a priori 18-item scales were very encouraging. Four of the five scales—Extraversion, Agreeableness, Conscientiousness, and Neuroticism—have shown excellent internal reliability in several studies (e.g., Grist & McCord, 2009; Hurt & Grist, 2010; Litty, 2007; McCord, 2006). These same scales correlate very well with similar constructs as measured by established temperament scales (e.g., Grist & McCord, 2010; Scheck & Grist, 2008). However, the Openness scale has thus far performed poorly in terms of both internal reliability and convergent validity.

The purpose of this project was to further refine the M5-PS item set and derive a new instrument that would maximize internal reliability for each scale, address serious problems with the Openness scale, improve external validity, and be shorter in length, making it more practical in applied settings such as preschool classrooms. In addition, cross-scale correlations were unacceptably high with the M5-PS-90, a problem experienced with other five-factor instruments; thus, particular effort was exerted at each step to emphasize discriminant validity of items and scales. We have accumulated sufficient data to divide the full sample randomly in half. In Phase I, with the first half of the combined data set, we conducted exploratory analyses on the 90 items to revise each personality factor scale. Once a tentative revised scale set was developed, we used confirmatory factor analysis on the second half of the data set to test its model fit. In Phase 2, data from a previous study were reanalyzed to compare the new 35-item scale to the original 90-item scale with regard to external validity.

PHASE 1—SCALE REFINEMENT

Method

Participants. There were 621 participants overall, randomly split into two groups, the first for use in an exploratory analyses to develop each five-factor scale, and the second for use in a confirmatory analysis. The first group consisted of 310 participants ranging in age from 2 to 6 years old ($M = 4.18$), half of whom were female and the other half male. The second group consisted of 311 participants ranging in age from 1 to 6 years old ($M = 4.17$), approximately half of whom were female and the other half male. All 621 participants were rated by their preschool teachers using the M5-PS-90.

Procedure.

Factor Refinement: We were guided by the creative approach employed recently by Tellegen et al. (2003) in the restructuring of the Basic Scales of the Minnesota Multiphasic Personality Inventory-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989). These researchers used factor analysis combined with other empirical and rational methods to develop relatively short “seed scales” that reflected the major distinctive core of each traditional clinical scale. They then computed correlations between the seed scale and each other item from the full 567-item set. They selected additional items for the final scale based on the size of the correlation but also content consistency, yielding longer scales that included the original seed scale plus other highly correlating items. The result was a new set of scales (the Restructured Clinical scales) with significantly

higher internal consistency and significantly lower intercorrelations. Similarly, our plan was to identify conceptually the major distinctive core of each of the five scales, to form small seed scales that represented these core concepts, and then to correlate all 90 items in our pool with each seed scale to produce final scales that are (relatively) unidimensional conceptually and optimized psychometrically.

We started with the 18 items in each factor on the M5-PS-90. In an effort to identify the major distinctive core of a given scale, we examined scale consistency data and item content. In particular, we noted the original facet association of the item on the adult scale as an indication of content focus, and we considered the premise that, among the 18 items on the original scale, the most frequently represented facet would likely reflect the major distinctive core for that scale. We also ensured that the content of each seed scale did not overlap with any other seed scale. This resulted in a five-item seed scale for Agreeableness composed of Items 18 (“Is easy to satisfy”), 27 (“Can’t stand confrontations”), 53 (“Yells at people”), 70 (“Insults people”), and 74 (“Gets back at others”). These items are all associated with the cooperation facet. The Conscientiousness seed scale consisted of four items: 4 (“Completes tasks successfully”), 22 (“Tries to excel at what they do”), 41 (“Does not understand things”), and 80 (“Doesn’t see the consequences of things”). These items are all on the self-efficacy facet. The Extraversion seed scale also had four items: 20 (“Radiates joy”), 28 (“Has a lot of fun”), 76 (“Laughs aloud”), and 83 (“Amuses his/her friends”). These items are on the cheerfulness facet. The Neuroticism seed scale had five items: 1 (“Worries about things”), 29 (“Is afraid of many things”), 40 (“Is relaxed most of the time”), 58 (“Gets stressed out easily”), and 79 (“Adapts easily to new situations”). These items all fall on the anxiety facet. Developing the Openness to Experience seed scale presented a new challenge. Adventurousness was the facet with the highest number of associated items (six). However, the content of these items clearly overlapped with Extraversion items, and to some extent with the negative poles of Conscientiousness and Neuroticism. In addition to the potential conceptual redundancy, we wanted to accommodate the possibility of including items associated with the imagination and artistic interests facets, which seem relevant in early childhood. Thus, we ended up selecting the six items for the O seed scale based on content, with three facets represented (imagination, artistic interests, and adventurousness): 2 (“Has a vivid imagination”), 13 (“Likes music”), 26 (“Dislikes changes”), 46 (“Does not like the idea of change”), 65 (“Is interested in many things”), and 73 (“Likes to begin new things”). Overall, the seed scales were considered to be representative of the major distinctive core of each personality factor.

Next, each of the 90 items was correlated with each seed scale (see Table 2). An item was retained on a scale (regardless of its original scale association) if: (a) its correlation with a seed scale was statistically significant and greater than its correlation with any other seed scale; (b) its correlation with the seed scale was greater than .5; and (c) its correlation with the second highest seed scale was at least .1 less in magnitude than the highest correlation. These rules were developed as a way to maximize discriminant validity as well as scale consistency as we formed the new scales. The new Agreeableness scale consisted of 15 items: 5, 8, 12, 18, 44, 53, 64, 67, 69, 70, 72, 74, 78, 87, and 88. Conscientiousness consisted of 10 items: 4, 19, 22, 37, 41, 51, 56, 71, 80, and 90. Extraversion consisted of five items: 20,

TABLE 2.—Item correlations with seed scales.

Item No.	A Seed	C Seed	E Seed	N Seed	O Seed	M5 Factor	New Factor
1	-.18	-.09	-.08	.63	-.24	N	N
2	-.09	.28	.39	-.09	.53	O	O
3	.19	.23	.28	-.48	.35	A	None
4	.27	.80	.39	-.27	.40	C	C
5	-.64	-.30	.00	.34	-.14	N	A
6	-.27	.26	.37	-.21	.38	E	None
7	-.20	-.05	-.03	.23	-.09	O	None
8	.54	.43	.02	-.15	.12	C	A
9	.16	-.16	-.32	.37	-.36	N	None
10	.27	.52	.54	-.47	.52	E	None
11	.19	.31	.42	-.47	.50	A	None
12	-.57	-.37	-.09	.43	-.23	N	A
13	.14	.24	.37	-.23	.48	O	None
14	-.38	-.15	.10	.41	-.12	O	None
15	.54	.58	.29	-.29	.34	C	None
16	-.08	.22	.51	-.31	.46	E	None
17	-.09	.07	-.01	-.19	.15	O	None
18	.65	.40	.27	-.44	.29	A	A
19	.13	.53	.35	-.30	.34	O	C
20	.31	.44	.74	-.41	.41	E	E
21	.45	.31	-.23	-.05	-.09	C	None
22	.21	.68	.45	-.37	.48	C	C
23	.48	.51	.26	-.25	.26	A	None
24	.00	-.12	-.22	.44	-.31	N	None
25	-.27	.04	.51	-.09	.30	E	E
26	.02	.19	.14	-.57	.58	O	None
27	.43	.01	-.07	.24	-.23	A	None
28	.09	.23	.68	-.32	.39	E	E
29	.07	-.19	-.29	.64	-.40	N	N
30	.11	.05	.08	.07	.01	O	None
31	.02	.14	.24	-.47	.33	A	None
32	.59	.57	.20	-.29	.24	A	None
33	-.11	-.18	-.25	.14	-.20	N	None
34	-.26	.31	.43	-.27	.43	E	None
35	-.22	-.12	-.04	.26	-.11	O	None
36	-.33	-.28	-.13	.07	-.06	E	None
37	.28	.63	.45	-.31	.44	C	C
38	-.15	.25	.52	-.35	.42	E	None
39	-.11	-.31	-.29	.68	-.52	N	N
40	-.35	-.37	-.41	.67	-.37	N	N
41	.12	.73	.36	-.31	.41	C	C
42	-.50	-.45	-.20	.61	-.40	N	N
43	.17	.37	.32	-.60	.54	E	None
44	.62	.15	-.08	-.15	-.01	A	A
45	.02	.28	.20	.18	-.02	C	None
46	.04	.17	.15	-.53	.60	O	None
47	-.23	.11	.48	-.23	.44	E	None
48	.00	.28	.53	-.56	.58	E	None
49	.17	.37	.47	-.49	.49	A	None
50	.30	.25	.08	.01	.03	C	None
51	.41	.55	.38	-.33	.41	A	C
52	-.18	-.18	-.21	-.05	-.05	O	None
53	.80	.30	-.09	-.32	.06	A	A
54	.08	.45	.49	-.23	.42	C	None
55	-.08	.51	.49	-.19	.43	O	None
56	.02	-.59	-.48	.20	-.44	A	C
57	-.37	-.54	-.33	.54	-.47	N	None
58	-.42	-.41	-.25	.75	-.45	N	N
59	.17	.38	.58	-.59	.58	E	None
60	.27	.40	.56	-.49	.56	E	None
61	.32	.31	.01	-.18	.11	C	None
62	-.41	.09	.30	-.02	.25	E	None
63	.39	.57	.49	-.31	.41	A	None
64	.60	.46	.23	-.23	.20	C	A
65	.11	.44	.58	-.31	.68	O	O
66	.08	.45	.50	-.34	.48	E	None
67	-.62	-.39	-.04	.46	-.27	N	A
68	-.28	-.14	-.11	.26	-.12	O	None
69	.66	.50	.08	-.24	.15	C	A

TABLE 2.—Item correlations with seed scales. (Continued)

Item No.	A Seed	C Seed	E Seed	N Seed	O Seed	M5 Factor	New Factor
70	.79	.35	.14	-.25	.10	A	A
71	.36	.59	.28	-.33	.38	C	C
72	-.72	-.30	.03	.36	-.12	N	A
73	.10	.47	.45	-.38	.72	O	O
74	.84	.25	-.05	-.21	.05	A	A
75	-.55	-.34	-.08	.56	-.30	N	None
76	-.08	.28	.75	-.15	.40	E	E
77	.05	.21	.21	.00	.08	A	None
78	.55	.43	-.07	-.23	.07	C	A
79	-.07	-.32	-.36	.71	-.58	N	N
80	.47	.66	.11	-.24	.15	C	C
81	.24	-.29	-.53	.30	-.44	N	None
82	.49	.00	-.32	.00	-.24	A	None
83	-.15	.33	.72	-.29	.42	E	E
84	.34	.57	.49	-.27	.35	O	None
85	-.32	-.24	-.11	.37	-.18	N	None
86	.55	.53	.40	-.37	.29	O	None
87	.56	.35	-.03	-.23	.07	A	A
88	.69	.32	-.02	-.31	.09	O	A
89	-.15	.36	.58	-.30	.52	C	None
90	.30	.69	.37	-.33	.43	C	C

25, 28, 76, and 83. Neuroticism consisted of seven items: 1, 29, 39, 40, 42, 58, and 79. No extra items met the inclusion criteria for the Openness to Experience scale; in fact, three of the actual seed scale items failed to meet our criteria, leaving three for the final O scale: 2, 65, and 73. Finally, Cronbach's alphas were computed for each of these new scales to remove items that did not contribute to each scale's internal consistency. We simultaneously considered the impact on cross-scale correlations resulting from item removal. That is, to achieve a final scale with an optimal combination of convergent and discriminant validity, with each item removal we tried to balance improvements in coefficient alpha with lowered cross-scale correlations. This final step resulted in the removal of Items 5, 12, 67, and 72 from Agreeableness, and Item 56 from Conscientiousness. Overall, 35 items were retained and 55 were discarded. The new instrument is labeled the M5-PS-35.

Confirmation of Factor Structure: Two confirmatory factor analyses (CFAs) were conducted to confirm the factor structure, using SAS (version 9.2). The first was on the M5-PS-90 and the second was on the M5-PS-35. The models were specified as oblique, allowing each factor to be correlated with each other factor. Oblique models do not require factors to be correlated and will produce orthogonal factors if that solution has the best simple structure (Fabrigar, Wegener, MacCallum, & Strahan, 1999). The models were fit using maximum likelihood estimation. The variance of each factor was constrained to 1.0, which is less restrictive than restraining unstandardized factor pattern coefficients. Each item was constrained to load on only one factor, which for the M5-PS-90 was the factor the item was originally associated with on the adult scale.

Goodness-of-fit indexes are "purely descriptive statistics" because little is known about their sampling distributions (Bentler, 1990). Many researchers therefore use different statistics and even interpret the same statistics differently. Hu and Bentler (1999) demonstrated through a series of simulations that using a combination of statistics is best because some statistics, such

as the standardized root mean squared residual (SRMR), are most sensitive to models with misspecified factor covariances, whereas others, such as the comparative fit index (CFI) and root mean square error of approximation (RMSEA), are most sensitive to models with misspecified factor loadings. Therefore, we used these three fit statistics in analyzing model fit. Hu and Bentler also suggested using combination rules, such as an SRMR less than .09, in combination with a CFI greater than or equal to .96 or a RMSEA of less than .06. Although these cutoff values would be ideal, it is highly unlikely for personality measures to meet these cutoffs (Hopwood & Donnellan, 2010; Marsh et al., 2010). Interpreting negative results from a CFA as a need to call into question the meaningfulness of the instrument should be avoided because many lower order scales have cross-loadings that might be minor but enough to increase model misfit for inventories with large bodies of evidence supporting their stability and predictive validity (Hopwood & Donnellan, 2010; Marsh et al., 2010). Thus, our intent is to use CFA statistics descriptively, to compare the M5-PS-35 to the M5-PS-90, in the context of CFA statistics typically reported for other five-factor personality instruments.

Because goodness-of-fit indexes approach unity (i.e., "perfect fit") by simply freeing up more parameters, we also included the parsimonious comparative fit index (PCFI; Mulaik, 2009; Mulaik et al., 1989). The PCFI combines information about goodness of fit with information about parsimony into a single index to compensate for artifactual increases in fit resulting from freeing up more parameters (Mulaik, 2009; Mulaik et al., 1989). Because some statistics, such as the Akaike information criterion (AIC), do not correct for model complexity as well as the PCFI and do not say anything about the fit of a model (i.e., a model in a set of models with the lowest AIC still might not fit the data very well, but the AIC alone does not tell us this), we used the PCFI to compare the M5-PS-35 with the M5-PS-90.

Results

Model fit was better for the further refined M5-PS-35 than for the M5-PS-90. This suggests that reducing the number of items ultimately resulted in better internal structure. The CFI for the M5-PS-35 was .743, much higher than the .416 for the M5-PS-90. Changes in the RMSEA and SRMR values were not as drastic, but still show improvement. The M5-PS-90 had an RMSEA of .093 and an SRMR of .157, whereas the M5-PS-35 had an RMSEA of .089 and an SRMR of .115. Finally, the M5-PS-35 had a PCFI of 0.688 and the M5-PS-90 had a PCFI of 0.406, indicating that the latent model of the M5-PS-35 is much better. In fact, it is not inconceivable to have parsimonious fit indexes in the .50s, although no acceptable values have been established (Mulaik et al., 1989). Overall the M5-PS-35 had better model fit than the M5-PS-90, with the SRMR and RMSEA coming closer to meeting their ideal cutoff values.

These results do not meet the cutoffs and are therefore less than ideal, but they are similar to fit indexes obtained in other well-known personality inventories with criterion-related validity such as the fifth edition of Cattell's 16PF, the Six-Factor Personality Questionnaire, the California Psychological Inventory, the HEXACO Personality Inventory, the Hogan Personality Inventory, the Multidimensional Personality Questionnaire, and the NEO PI-R, all of which had CFI values ranging from .61

TABLE 3.—Confirmatory factor analysis structure coefficients for the 35-item scale.

Factor	Item No.	Structure Coefficient	Factor	Item No.	Structure Coefficient
Agreeableness	8	.79	Extraversion	20	.65
	18	.63		25	.52
	44	.56		28	.67
	53	.73		76	.61
	64	.61		83	.64
	69	.80	Neuroticism	1	.26
	70	.69		29	.50
	74	.65		39	.61
	78	.63		40	.62
	87	.75		42	.72
	88	.72		58	.77
Conscientiousness	4	.75		79	.55
	19	.67	Openness to Experience	2	.57
	22	.77		65	.71
	37	.79		73	.75
	41	.54			
	51	.58			
	71	.47			
	80	.39			
	90	.73			

to .79 and RMSEA values ranging from .09 to .13 (Hopwood & Donnellan, 2010). The M5-PS-35's results were also similar to the adult M5-50 (Socha, Cooper, & McCord, 2010) and the 60-item NEO-FFI (Marsh et al., 2010). Each item loaded significantly on its factor for the M5-PS-35 (see Table 3 for CFA structure coefficients on the target factors; coefficients on the nontarget factors were set to zero). This indicates that each item has a significant contribution on its personality factor and is further evidence of construct validity.

Table 4 presents intercorrelations among the five scales for both the 90-item and 35-item versions, and for comparison purposes we included cross-scale correlations from Goldberg's (2001) reanalysis of the Digman data, and for the short form of the ICID. As can be seen in Table 4, cross-scale correlations are far from trivial. Although the factors of the FFT have been described in abstract as "relatively orthogonal," in terms of actual measurement there are typically some significant correlations among the scales. Thus, a realistic goal is to achieve correlations that are at least consistent with, if not lower than, those found with other five-factor instruments. For example, as shown in Table 4, the strongest three correlation coefficients for the M5-PS-35 are .60 (C/O), .57 (E/O), and .54 (A/C). In comparison, the strongest three for the M5-PS-90 were .94 (A/C), .89 (C/O), and .84 (A/O). The top three for the short form of the ICID were -.84 (A/N), .74 (E/O), and -.61 (C/N). Not shown are the even higher correlations for the longer, original version of the ICID, with the top three being -.87 (A/N), .70 (C/O), and -.65 (C/N). Although the problem of cross-scale correlations seems particularly challenging for those of us working with child samples, it should be noted that the three strongest cross-scale correlations for the NEO PI-R, reported in the technical manual using the normative sample, were .53 (C/N), .40 (E/O), and .24 (A/C; Costa & McCrae, 1992b). Of most direct relevance to the M5-PS-35 are the correlations reported by Goldberg (2001) in his reanalysis of the Digman data. As shown in Table 4, the strongest three correlations were .41

TABLE 4.—Cross-scale correlations for the M5-PS-35, with comparison data.

		E	A	C	N	O
M5-PS-35 (<i>n</i> = 311)	E	—	.07	.43	-.36	.57
	A		—	.54	-.47	.20
	C			—	-.41	.60
	N				—	-.36
	O					—
M5-PS-90 (<i>n</i> = 311)	E	—	.29	.30	-.37	.68
	A		—	.94	-.77	.84
	C			—	-.57	.89
	N				—	-.65
	O					—
ICID (<i>n</i> = 903)	E	—	.32	.31	-.32	.74
	A		—	.56	-.84	.42
	C			—	-.61	.31
	N				—	.40
	O					—
Digman ^a	E	—	-.30	.10	-.10	.38
	A		—	.41	-.35	.02
	C			—	-.36	.39
	N				—	.19
	O					—

Note. E = Extraversion; A = Agreeableness; C = Conscientiousness; N = Neuroticism; O = Openness to Experience; ICID = Inventory of Children's Individual Differences. ICID data are based on normative sample for the most recent short version (Deal et al., 2007) provided by Deal, personal communication, April 1, 2011. Digman caption refers to Goldberg's (2001) analyses of Digman's data. With the Digman data, Emotional Stability is presented here as the N factor with reverse sign on correlation coefficient, to facilitate comparison.

^aDigman data are based on 2,572 children overall, but the data reported are average correlations across the six separate samples, which varied in size from 502 to 885 children.

(A/C), .39 (C/O), and .38 (E/O). The cross-scale correlations for the M5-PS-35 are somewhat higher than those for the Digman data but are far lower than the correlations found with the original 90-item version and with both versions of the ICID. Digman (1997) proposed a higher order structure for five-factor theory, which makes predictions about patterns of cross-scale correlations. Briefly, he proposed an "alpha" factor consisting of A, C, and N, with a second "beta" factor consisting of E and O. Correlations between scales within each factor should be higher than with scales in the other factor. Goldberg (2001) concluded that, across all samples, the Digman data are "not inconsistent" with this hypothesis. Indeed, if we use .30 as a cut point and count how many of the 10 correlation coefficients are consistent with the hypothesis, the Digman data score 8 out of 10. The M5-PS-35 scores 6 of 10. The ICID is not directly comparable because all 10 correlation coefficients exceed .30.

Cronbach's alphas and average interitem correlations were computed using SAS (version 9.2) on the confirmatory analysis group of our sample to ensure that reliability did not decrease. Reliability was similar between the personality factors of both instruments, with the exception of Openness to Experience, which had a Cronbach's alpha of .53 on the M5-PS-90, but increased to .71 on the M5-PS-35. This is a major increase in reliability for Openness to Experience. Coefficient alphas as well as mean interitem correlations are presented in Table 5.

PHASE 2—VALIDATION STUDY

Grist and McCord (2010) compared preschool teacher ratings of children in their classrooms on the M5-PS-90 and the three temperament scales of the Child Behavior Questionnaire. We reanalyzed their data using the 35-item set.

TABLE 5.—Cronbach's alphas and average interitem correlations.

	A		C		E		N		O	
M5-PS-	35	90	35	90	35	90	35	90	35	90
Cronbach's alpha	.90	.79	.87	.89	.77	.88	.79	.89	.71	.53
Average interitem correlation	.45	.19	.42	.31	.40	.30	.35	.31	.45	.06

Note. A = Agreeableness; C = Conscientiousness; E = Extraversion; N = Neuroticism; O = Openness to Experience.

Method

Participants. Participants were 122 children 3 to 4 years of age in regular preschool classrooms in North Carolina. Parental permission was obtained for teachers to rate the children.

Measures.

M5-PS-90: This instrument was described thoroughly earlier. The full 90-item instrument was completed by teachers. We extracted the new 35-item set from the full item set for purposes of comparison.

Rothbart Childhood Behavior Questionnaire—Very Short Form (CBQ): Preschool teachers also completed the CBQ, a temperament measure designed to assess three dimensions: Surgency, Negative Affect, and Effortful Control. These scales have consistently emerged during factor analysis from the standard form of the CBQ (Putnam & Rothbart, 2006; Rothbart, Ahadi, Hersey, & Fisher, 2001). Putnam and Rothbart (2006) reported acceptable internal reliability for the CBQ, with alphas of .75 for Surgency, .72 for Negative Affect, and .74 for Effortful Control. (Alpha values for this study were .82 for Surgency, .66 for Negative Affect, and .85 for Effortful Control.)

Procedure. Materials were delivered in packets to the preschool teachers who agreed to participate in the study. The two questionnaires were arranged in counterbalanced order. Teachers were given approximately 2 weeks to complete ratings of all children in their classrooms.

Results

We reanalyzed data using both the M5-PS-90 and the M5-PS-35 scale scores. Results are presented in Table 6. The most clear-cut hypotheses were that the personality factor Extraversion should correlate positively with the temperament factor Surgency; Conscientiousness should correlate positively with Effortful Control; and Neuroticism should correlate positively with Negative Affect. Overall, results of the two versions

TABLE 6.—Comparisons of M5-PS-35 versus M5-PS-90 correlations with Rothbart temperament scales.

	E		A		C		N		O	
M5-PS-	35	90	35	90	35	90	35	90	35	90
Surgency	.61	.59	-.44	-.32	-.17	-.36	-.26	-.13	.15	.01
Negative affect	-.13	-.11	-.37	-.44	-.07	.04	.62	.65	.01	-.08
Effortful control	-.02	.03	.21	.16	.50	.54	.14	.04	.26	.20

Note. Bold font indicates $p < .01$ and underline indicates $p < .05$. E = Extraversion; A = Agreeableness; C = Conscientiousness; N = Neuroticism; O = Openness to Experience.

of the M5-PS are quite similar. Using the Fisher z score transformation procedure to compare correlation coefficients, there were no cases in which the correlation coefficients produced by the two different forms of the M5-PS differed significantly from each other. Primary hypotheses were strongly supported (for both forms), with substantial correlations between E and Surgency, N and Negative Affect, and C and Effortful Control. Because the correlations did not differ significantly from each other, we cannot state that the M5-PS-35 exhibited greater validity than the M5-PS-90; however, the fact that a new instrument of less than half the length of the old, with improved internal psychometric characteristics, is able to perform equally well as the longer version suggests that the M5-PS-35 is a potentially useful contribution to the field.

GENERAL DISCUSSION

This study was designed to derive a short, reliable, and valid instrument for measuring the FFT in preschool children. Ninety items were shortened down to 35 through content focusing and removing unreliable items from each factor, resulting in increases in overall reliability and construct validity. The potential for major gains in discriminant validity is indicated by significantly lower cross-scale correlations. Comparisons with external criteria yielded results that were generally very similar between the original 90-item form and the new 35-item version.

There is certainly room for improvement. The first area that needs further refinement is the Openness to Experience factor. This factor was initially the most unreliable out of the five personality factors and was trimmed down to three items. The increase in Cronbach's alpha for the O factor from an unacceptable .53 on the 90-item version to .71 on the new 35-item represents a major improvement. Even so, a new pool of items should be derived and tested to increase the reliability and validity of measuring this factor in preschool children.

With regard to the degree of intercorrelations among scales, the 35-item scale resulted in clear improvement. The overall average magnitude (ignoring sign) of the 10 cross-scale correlations for the new 35-item instrument was .40, which is far better than the .63 for the 90-item version and compares favorably to the .48 for the ICID. In the 35-item instrument, the highest correlation was .60, between C and O; this was also the highest correlation in the Digman data, although the correlation coefficient was just .39. In our 90-item version, 6 of the 10 correlations exceeded .60, and in the ICID short form 3 of the correlations exceed this level. In our new 35-item version, all intercorrelations are .60 or lower.

Our CFA results were below the recommended cutoffs for formal model testing. The 35 resulted in clear improvement compared to the 90, yet it is still below ideal test values. First, we should stress the fact that CFA was used as a descriptive comparative technique, helping to improve the instrument, rather than as a pure test of model fit. Although there are some instances in which a five-factor instrument does indeed meet, or approach, ideal CFA statistics (e.g., Socha et al., 2010), in most cases the recommended cutoffs for theory testing are not met when testing five-factor and other personality measurement instruments (Marsh et al., 2010). Marsh et al. (2010) addressed this issue very specifically, noting that exploratory factor analysis was used extensively in the development of the FFT, but CFA has failed to provide clear support for the model. Marsh et al.

echoed the views of others (e.g., Costa & McCrae, 1992b, 1995; McCrae & Costa, 1997) that the primary problems are not with the FFT itself but with CFA as a model-testing approach. Our own view is that the CFA techniques can provide useful descriptive information for comparing solutions in the current context, although resulting models should not be evaluated strictly by recommended cutoffs.

We had originally hoped to retain a sufficient number of items and within-scale content heterogeneity that narrower facet traits could be discovered, through future research efforts. The 15 narrow-band traits measured by the ICID, in addition to the five broad domains, seem potentially very useful. Maintaining this multifacet content approach proved to be untenable, however, as reflected in lowered alpha coefficients and unacceptable increases in cross-scale correlations. For example, items reflecting "anger" tend to appear both on the A scale (in the negative direction) and on the N scale. Items reflecting "adventurousness" could appear on E or O. Cooperative and compliant items could be A as well as C. Thus, we had to make choices about the major distinctive core concept for each of the five domains. One undesired outcome of this approach was that we have likely constrained the possibility of discovering facets, or meaningful subscales, within this current item set.

Despite these limitations, the M5-PS-35 exhibits good internal consistency for a research tool, with acceptable construct validity. Currently available data suggest that the shorter version exhibits at least equivalent convergent validity and the potential for significant improvements in discriminant validity, although clearly more work is needed in these areas. In the spirit of Goldberg's IPIP "collaboratory," we offer this as a freely accessible and modifiable public-domain instrument for researchers interested in studying the Big Five in preschool children.

REFERENCES

- Abe, J. A. A., & Izard, C. E. (1999). A longitudinal study of emotion expression and personality relations in early development. *Journal of Personality and Social Psychology*, 77, 566-577.
- Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. Burlington: University of Vermont, Department of Psychiatry.
- Asendorpf, J. B., & Denissen, J. J. A. (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—Journal of Developmental Psychology*, 52, 486-513.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Block, J. (1995). A contrarian view of the five-factor approach to personality. *Psychological Bulletin*, 117, 187-215.
- Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality development*. New York, NY: Wiley.
- Buss, D. M. (1991). Evolutionary personality psychology. *Annual Review of Psychology*, 42, 459-491.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI-2: Minnesota Multiphasic Personality Inventory-2: Manual for administration and scoring*. Minneapolis: University of Minnesota Press.
- Costa, P. T., & McCrae, R. R. (1992a). Four ways five factors are basic. *Personality and Individual Differences*, 13, 653-665.
- Costa, P. T., & McCrae, R. R. (1992b). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1994). Stability and change in personality from adolescence through adulthood. In C. F. Halverson, G. A. Kohnstamm, & R.

- P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 139–150). Hillsdale, NJ: Erlbaum.
- Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, 64, 21–50.
- De Clercq, B., & De Fruyt, F. (2003). Personality disorder symptomatology in adolescence: A Five-factor model perspective. *Journal of Personality Disorders*, 17, 269–292.
- De Fruyt, F., Mervielde, I., Hoekstra, H. A., & Rolland, J.-P. (2000). Assessing adolescents' personality with the NEO PI–R. *Assessment*, 7, 329–345.
- De Pauw, S. S. W., Mervielde, I., & Van Leeuwen, K. G. (2009). How are traits related to problem behavior in preschoolers? Similarities and contrasts between temperament and personality. *Journal of Abnormal Child Psychology*, 37, 309–325.
- Deal, J. E., Halverson, C. F., Martin, R. P., Victor, J., & Baker, S. (2007). The Inventory of Children's Individual Differences: Development and validation of a short version. *Journal of Personality Assessment*, 89, 162–166.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 41, pp. 417–440). Palo Alto, CA: Annual Reviews.
- Digman, J. M. (1997). Higher order factors of the Big Five. *Journal of Personality and Social Psychology*, 73, 1246–1256.
- Digman, J., & Shmelyov, A. (1996). The structure of temperament and personality in Russian children. *Journal of Personality and Social Psychology*, 7, 341–351.
- Eysenck, S. B. G., Makaremi, A., & Barrett, P. T. (1994). A cross-cultural study of personality-Iranian and English children. *Personality and Individual Differences*, 16, 203–210.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48, 26–34.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg, The Netherlands: Tilburg University Press.
- Goldberg, L. R. (2001). Analyses of Digman's child-personality data: Derivation of Big-Five factors scores from each of six samples. *Journal of Personality*, 69, 709–743.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1), 84–96.
- Gosling, S. D., & John, O. P. (1999). Personality dimensions in nonhuman animals: A cross-species review. *Current Directions in Psychological Science*, 8, 69–75.
- Grist, C. L., & McCord, D. M. (2009, February). *A personality scale for preschoolers: Preliminary psychometrics*. Poster session presented at the annual meeting of the Southeastern Psychological Association, New Orleans, LA.
- Grist, C. L., & McCord, D. M. (2010). Individual differences in preschool children: Temperament or personality? *Infant and Child Development*, 19, 264–274.
- Halverson, C. F., Havill, V. L., Deal, J., Baker, S. R., Victor, J. B., 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, 995–1026.
- Hogan, R., & Hogan, J. (1992) *Hogan personality inventory manual*. Tulsa, OK: Hogan Assessment Systems.
- Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review*, 14, 332–346.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- Hurt, A., & Grist, C. L. (2010, March). *Personality trait differences between American and Jamaican preschool children*. Poster session presented at the annual meeting of the Southeastern Psychological Association, Chattanooga, TN.
- International Personality Item Pool. (2001). *A scientific collaboratory for the development of advanced measures of personality traits and other individual differences*. Retrieved from <http://ipip.ori.org/>
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York, NY: Guilford.
- Kohnstamm, G. A., Halverson, C. F., Jr., Mervielde, I., & Havill, V. L. (1998). *Parental descriptions of child personality: Developmental antecedents of the Big Five*. London, England: Erlbaum.
- Litty, C. G. (2007). *Temperament and personality in preschoolers: Are the concepts same or different* (Doctoral dissertation). Retrieved from <http://etd.utk.edu/2007/LittyCathy.pdf>
- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., & Trautwein, U. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*, 22, 471–491.
- McCord, D. M. (2006, March). Evidence for the Five factor model in young children. In C. G. Litty (Chair), *Individual differences in preschool children: Temperament or personality?* Symposium conducted at the annual meeting of the Southeastern Psychological Association, Atlanta, GA.
- McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52, 509–516.
- Measelle, J. R., John, O. P., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (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, 90–106.
- Mervielde, I., De Clercq, B., De Fruyt, F., & Van Leeuwen, K. (2005). Temperament, personality and development psychopathology as childhood antecedents of personality disorders. *Journal of Personality Disorders*, 19, 171–201.
- Mervielde, I., & De Fruyt, F. (2002). Assessing children's traits with the Hierarchical Personality Inventory for Children. In B. DeRaad & M. Perugini (Eds.), *Big five assessment* (pp. 129–146). Seattle, WA: Hogrefe & Huber.
- Mulaik, S. A. (2009). *Linear causal modeling with structural equations*. Boca Raton, FL: Chapman & Hall/CRC.
- Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, 105, 430–445.
- Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, 61, 622–631.
- Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment*, 87, 102–112.
- Rothbart, M. K., Ahadi, S. A., Hersey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development*, 72, 1394–1408.
- Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (pp. 37–86). Hillsdale, NJ: Erlbaum.
- SAS (Version 9.2) [Computer software]. Cary, NC: SAS Institute.
- Scheck, A., & Grist, C. L. (2008, March). *Personality and temperament: A comparison in a preschool population*. Poster session presented at the annual meeting of the Southeastern Psychological Association, Charlotte, NC.
- Socha, A., Cooper, C. A., & McCord, D. M. (2010). Confirmatory factor analysis of the M5–50: An implementation of the IPIP item set. *Psychological Assessment*, 22(1), 43–49.
- Tellegen, A., Ben-Porath, Y. S., McNulty, J. L., Arbisi, P. A., Graham, J. R., & Kaemmer, B. (2003). *MMPI–2 Restructured Clinical (RC) scales: Development, validation, and interpretation*. Minneapolis: University of Minnesota Press.
- Thomas, A., & Chess, S. (1977). *Temperament and development*. New York, NY: Brunner/Mazel.