Validation of the School-age Assessment of Attachment in a short-term longitudinal study

Patricia Crittenden¹, Katrina Robson² and Alison Tooby²
¹Family Relations Institute, USA
²Cumbria Partnership NHS Foundation Trust, UK

Abstract
This study tested the validity of the School-aged Assessment of Attachment (SAA) in terms of matches from the well-validated Preschool Assessment of Attachment (PAA) to the SAA 6 months later. It also addressed validity in terms of mental health services and measures of stress, depression and anxiety.

Hypotheses: Children’s SAA classifications were predicted to match their 6 months’ previous PAA classifications and indicators of maternal, child and family stress. The study used a two-group comparative design, involving normative and clinical children and their mothers.

Method: The participants were 50 children between 5.5 and 5.9 years of age. Each child participated with his or her mother in a PAA, and then 6 months later each child responded to the SAA story cards as well as self-report assessments of stress, anxiety and depression.

Results: Concordance of A, B, C and A/C attachment classifications was found between the PAA and SAA in 34 of 48 children. There was a strong relation between referral status (clinical or normative) and both PAA and SAA attachment classifications. In every non-matching case, a normative child had an attachment classification indicative of risk, indicating that the direction of errors was false positives as opposed to false negatives.

Conclusions: This evidence supports the validity and clinical utility of the SAA.

Keywords
Attachment, DMM, PAA, SAA, school-age children

As the study of attachment and its clinical applications has expanded from its base in infancy through the preschool years and into adulthood, interest has focused on the school years. Up to now, there has not been clarity about the development of attachment in the school years, nor ways to assess it. This article is one in a planned set of studies to validate one approach to assessing attachment in children from 6 years of age to puberty, the School-aged Assessment of Attachment.

Corresponding author:
Patricia Crittenden, Family Relations Institute, 9481 SW 47 St., Miami, FL 33176, USA.
Email: pmcrittenden@gmail.com
(SAA; Crittenden, 2009). Because many assessments of attachment have been developed, but few have shown reliability, validity and utility, we begin by discussing validity in psychological testing, and then move to discussing how to develop an assessment of attachment for this age group. Following that, we turn to development in the school years and then to the SAA itself. The focus throughout is the transition from enacted infant strategies to linguistically represented strategies in adulthood, with the school years falling at the transition point.

Validity in psychological testing

Although many types of validity have been considered over the last century, by late 20th century, (a) ‘construct validity’ was considered to underlay all types of validity (Chronbach, 1971; Chronbach & Meehl, 1955; Messick, 2000), (b) validation was considered to be an on-going process (as opposed to establishing validity once and for all; Kane, 2013) and (c) the validity of specific assessments was understood to lie equally in theoretical coherence, empirical measurement and the specific applications of the tools. Applications referred to the kind of decisions that would be made about individuals based on their responses using the assessment (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). In particular, the consequences of the applications for individuals and professionals took on importance (Messick, 2000). For professionals, this meant whether use of the tool required special skills and authorizations (Newton & Shaw, 2014).

The result is that validation of assessment tools is understood to include an extended programme of studies (rather than a single study), the need to specify the constructs underlying the tool before testing the tool and the need to specify the interpretation of the results, including comparing it with competing assessments (Kane, 2013). These ideas are particularly attuned to written educational and psychological ‘tests’. Attachment, however, is not ‘tested’ and, instead, its quality or interpersonal patterning is assessed through structured observations. A tradition of validating assessments of attachment has been developed and applied to new assessments of attachment as they are offered to the professional community.

Assessing attachment after infancy

This review is limited to those assessments that contributed directly to the development of the SAA (review of the many other assessments in use exceeds the purview of this article).

The Strange Situation Procedure for infants

The gold standard for assessing attachment is Ainsworth’s Strange Situation Procedure (SSP, Ainsworth, Blehar, Waters, & Wall, 1978); its success may lie in how it was produced. The SSP was grounded in both (1) theory about separation and maternal-sensitive responsiveness and (2) a well-known sample that Ainsworth and her colleagues had seen at home for a year. The laboratory context and eight episodes (3) structured the responses of infants and their mothers around a gradient of (4) age-salient and (5) real threat (separation from the mother), to elicit children’s self-protective attachment strategies. The strategies were enacted by infants, thus (6) fitting infants’ cognitive and communicative abilities. Following the observations in the laboratory, Ainsworth (7) extracted a set of response categories (the classificatory system) that fit her normative sample and (8) drew up coding guidelines (the classificatory method) for categorizing the protocols. She then (9) replicated the SSP on a similar sample. Other investigators tested the (10) longitudinal validity of the SSP and (11) expanded the classificatory system to less normative samples (A/C: Crittenden,
1985; disorganized: Main & Solomon, 1986). From its implementation in the mid-1960s and first publication in 1969 to general acceptance, in the mid-1980s, the validation of the SSP took about two decades.

Building on this accomplishment, other assessments were developed for older ages. Some were accepted as valid whereas others were not. The reasons are relevant to the development of new assessments of attachment, particularly the SAA.

**The Adult Attachment Interview**

Main and her colleagues repeated Ainsworth’s process, with one exception, to develop the Adult Attachment Interview (AAI; George, Kaplan, & Main, 1985/1996; Main, Kaplan, & Cassidy, 1985). They began with Bowlby’s theory about adults’ information processing about threat (Bowlby, 1980), used a sample of mothers previously seen with infants in the SSP, constructed a semi-structured interview (like the eight episodes of the SSP), introduced an age-salient and real threat (thinking about threats in one’s childhood in the context of being a parent), tapped the mothers’ attachment strategies with verbal representations (rather than enactment), thus fitting adults’ cognitive/linguistic abilities. The exception was applying the infant classificatory system to adults. Other investigators replicated and extended the original findings and expanded the classificatory system to Cannot Classify (Hesse, 1996).

The limitations of the AAI are the difficulty of reaching reliability and constraining adult outcome categories to those used by infants. Later approaches to coding and classifying AAI transcripts expanded the classificatory system to capture adult strategies not available to infants, expanded the discourse analysis to include markers and distinctions not addressed in Main’s classificatory method and used the enacted process between speaker and interviewer as part of the classificatory process (Crittenden & Landini, 2011). This resulted in greater clinical validity and utility, but lower reliability, with greater complexity leading to fewer coders reaching reliability. The AAI has taken more than two decades to validate for risk populations.

When the AAI is applied to adolescents, it loses much of its power to differentiate individuals who differ in adaptiveness. Given adolescents’ less complete cortical development and different age-salient concerns (such as self-identity and dating as opposed to raising children), the AAI may be developmentally inappropriate for adolescents (Black, Jaeger, McCartney & Crittenden, 2000).

**Attachment in preschool-aged children**

The first approach to developing an assessment for preschool-aged children was somewhat different. It began by using the SSP, with enactment of children’s strategies, but expanded outcome categories and changed coding guidelines to fit theory and observation of 2- to 5-year-old children. Rather than being based on a known sample, a set of mostly normative SSPs was accumulated by a multi-university committee (MAC; Cassidy, Marvin & The MacArthur Attachment Working Group of the John D. and Catherine T. MacArthur Network on the Transition from Infancy to Early Childhood, 1992). The infant categories were retained and new categories were defined as analogous to, or developmentally evolved from, infant disorganization. The committee emphasized coder reliability, with validity to be tested in the studies that would follow. The results were encouraging regarding reliability, but less strong than the SSP regarding validity, especially in clinical samples. Crittenden applied the MAC to her Virginia sample and found confusion of classification between normative and maltreated children (Crittenden, Claussen, & Kozlowska, 2007). Using this sample, Crittenden developed both a set of expanded classifications (A3-4 and C3-4) and new classificatory criteria, then trialled the resulting Preschool Assessment of Attachment (PAA) on a...
Texas sample, using that sample to train coders to recode the Virginia protocols with the PAA. Comparison of the MAC and PAA classifications on the Virginia sample indicated greater clinical validity for the PAA (Crittenden et al., 2007). This was replicated on the National Institute of Child Health and Development Study of Early Child Care and Youth Development sample (Spieker & Crittenden, 2009).

The comparative findings are important because they clarify the process needed to generate a valid assessment of attachment that also has clinical utility. Specifically, a sample of children who are known to the researchers and who vary in adaptation facilitates building validity into the coding guidelines. In addition, consideration of developmental change in organizational processes promotes discovery of useful outcome categories. These two omissions limited the clinical utility of the MAC as compared to the PAA. Unlike the SSP, it took only a decade to test the validity of the MAC and PAA.

**Attachment in 4- to 5-year-olds**

The development of the Narrative Story-stems (NSS) used yet another process (Bretherton, Ridgeway, & Cassidy, 1990). The NSS attempted to capitalize on 4- to 5-year-olds’ emerging cognitive/linguistic representational skills by asking them to act out stories with figurines while describing in words what was happening. The stories were prompted by a narrative ‘stem’ that addressed common problems that children experience (e.g. spilling juice). Coding guidelines were later developed, after many investigators (including the first author of this article) had used the NSS procedure. Because the results were often unclear, most studies were not published. Part of the problem, early on, was lack of a formal coding method. Other problems were more enduring. The story topics are not very threatening; the issue is whether they activate children’s self-protective strategies. If they do, then children’s exploratory behaviour should decrease. Instead, many children become quite playful, moving around the room, sometimes throwing the figurines and making accompanying sound effects. Such enactment, combined with linguistically represented information, creates coding difficulties. This complexity is increased by the relatively unstructured NSS procedure. Finally, having 4- to 5-year-olds hold, move, and make sounds for the figurines presents developmental problems regarding enactment versus representation. The NSS assumes that children can differentiate themselves from figurines. That may anticipate skills that do not develop until children gain access to ‘source memory’ at about 7 years of age (Schacter, 1996). Before that, children confuse different sources of information (e.g. the child’s experience, the child’s imagination and a story told by someone else). Confusion is increased when children tell an imaginary story while holding a figurine and acting out the figurine’s story. In almost 25 years, the NSS procedure has not yet been adequately validated.

Attachment in the school years should be viewed against this backdrop of assessments working forward from infancy and backward from adulthood.

**Assessing attachment in the school years**

*The development of attachment in the school years*

One crucial feature of attachment in the school years is a shift from parents functioning as children’s sole attachment figures to the addition of best friends as attachment figures and the need to coordinate parent, best friend and peer-group relationships. A second major change is the greater emphasis upon exploration away from parents – in the neighbourhood, school, organized activities and even the fantasy domains of television, electronic games and books. These two changes signal
a change in the dangers to which children are exposed. Instead of separation from the parent being central, other dangers such as peer rejection, family discord and break-up, and injury or illness become prominent. These dangers require more elaborate protective strategies than those used by preschoolers. Building on the classification system of the PAA, two new Type C strategies were defined (C5: coercively punitive and C6: coercively seductive). Both depend upon the ability to deceive using ‘false cognition’ (see Figure 1).

Psychologically, school-aged children think in concrete ways; they can use pictorial representations of dangerous situations to elicit linguistic representations. Because cognitive/linguistic skills are being learned, children’s representational capacity changes from limited linguistic representation at 6 years of age to sophisticated representation by puberty. Functionally, school-aged children can generate semantic generalizations about their own experience, tell episodes about their experience and use connotative language to express feelings. Their generalizations are limited, however, by a tendency to dichotomize complex situations (e.g. a parent or best friend who has mixed good and bad qualities is conceptualized as all good or all bad) and to change from day to day (e.g. children ‘loving’ a parent who does what they want and ‘hating’ a parent who punishes them). School-aged children also find it difficult to place episodic experiences in temporal sequence. On the other hand, school-aged children, unlike their former preschool-aged selves, are able to keep information secret. Secrets indicate a theory of mind and form the basis for intrapersonal privacy, interpersonal confidence and deception of other people. In addition, school-aged children construct make-believe, imagined or fantasy stories and use these to explore beyond their direct experience, express their wishes and talk about troublesome experiences that they cannot face directly. This makes it difficult to interpret imagined stories. Nevertheless, such representations are clinically important.

Crucially, school-aged children cannot accurately represent family situations that they do not observe completely or which result from hidden conditions. For example, marital discord, partner violence, separation, divorce and even having unmarried parents affect children strongly, but are incompletely understood by children and, as a consequence, can only be represented incompletely by them. Unfortunately, the more inexplicit a threat is, the greater is its threat to the child (because it cannot be addressed directly by the child) and the less clear is its representation by the child (thus making it difficult for others to discern what the child perceives as threatening). In such cases, imaged and somatic forms of representation may predominate. That is, when
threat cannot be represented directly, nor enacted behaviourally, it may be inhibited (repressed), transformed into fantasy or embodied. To be clinically useful, beyond the symptoms that bring children to clinical attention, an assessment should elicit children’s representational process in ways that give meaning to the presenting behaviour or signal a need to attend to distress that has not yet elicited adult concern.

The SAA

The SAA is based on developmental theory about attachment in the school years (see above, Crittenden, 2008, 2015) and the clinical applications of attachment theory to children (Crittenden, Dallos, Landini, & Kozlowska, 2014). The SAA uses picture cards with a semantic label and asks for both imagined stories about the child on the card and recalled episodes about the self in similar circumstances. It is an attempt to bridge the gap from enacted assessment in the SSP and PAA to linguistic representation in the AAI.

The seven SAA picture cards present a gradient of dangers experienced by school-aged children (from going out alone through rejection by friends, family separation and running away to mother going to hospital). The seven dangers were chosen because each could be associated with a wide range of similar events. For example, the ‘father going away’ card could reflect the father going to work, the father going on a trip, the father leaving after an argument, the absence of a father in the family or even the mother abandoning the family. The purpose of the pictures is to focus the child on concrete situations, with the fantasy story and recalled episode replacing enactment, but each in a different way. The fantasy stories allow children to convey concerns that might not have happened or to transfer real events into fantasy, thus distancing the events from the self. The recalled episodes provide evidence of actual events, as reconstructed by the child; that is, it is not assumed that recalled events are told with veridical accuracy or as other participants experienced them. Semi-structured follow-up questions ensure full coverage of each threat.

The outcome classifications of the SAA consist of three parts. The protective strategies include Ainsworth’s A1-2, B and C1-2 patterns of attachment, plus strategies from the Dynamic-Maturational Model of Attachment and Adaptation (DMM): A/C, the PAA A3-4, C3-4 patterns and two new ‘coercive’ Type C strategies (C5-6; Crittenden, 1985, 1994) (see Figure 1). In addition, the strategy can be in a ‘depressed’ form, that is, non-strategic with low arousal and low expectation of self-efficacy. Finally, psychological trauma from exposure to danger may be noted. Because of the breadth and non-specificity of the stimulus cards, coders’ agreement is only for the assignment of some psychological trauma. For clinical application, further information is needed.

The purpose of the gradient of increasing threat (as in the AAI, SSP and PAA) is to elicit the child’s protective attachment strategy; in the SAA, the strategy is the psychological process for managing information about threat. Coding is based on transcribed discourse; the discourse markers were derived from the DMM-AAI markers (Crittenden & Landini, 2011), but limited to those observed in early exploratory SAAs. In addition, the content of the stories is coded in terms of temporal causation, affect, perspective taking and generation of alternative solutions if the event were to occur again. As with other assessments of attachment, the SAA requires trained and reliable coders who use the manual to identify a child’s protective attachment strategy. The quality of the child’s discourse indicates the degree of arousal when thinking about such events in general and specific lapses suggest aspects of the event that distress the child. Once the discourse is annotated, transforming the detail into the child’s concrete linguistic representations requires attention to both the content of children’s stories and also the discourse with which the
story is told. This is different from the history component of an AAI which can include a very
dangerous event, but be considered secure if the discourse is coherent. Consistent with concrete
mental functioning, in the SAA, the events themselves carry weight, separate from the coherence
of the discourse. As is true of all observational assessments of attachment (from the SSP to the
AAI), the coder’s classification is inferential and, in childhood, is limited by the developmental
competencies of children.

Validation of the SAA

As with all assessments, the SAA needs demonstrated reliability among coders, validity of out-
comes and utility of application. Utility is crucial. A number of instruments have demonstrated
statistically significant differences between normative and risk groups, but with high proportions
of false positives and false negatives. This is true for the MAC, the Child Attachment Interview
(Shmueli-Goetz, Target, Fonagy, & Datta, 2008) and for self-report measures. The self-report
instruments are particularly problematic because they are easily affected by the cognitive limita-
tions of school-aged children.

In addition, differentiation within the risk group is important. If a tool merely duplicates clinical
information regarding risk, it will have limited utility. The tool should add to what is already
known. To do this, the tool must have a range of outcomes that reflect differences in symptoms and
adaptation among risk children (as opposed to placing most risk children in a few general catego-
ries such as insecure, insecure other, unresolved, disorganized or cannot classify). Because the
SAA has a wide range and gradient of possible outcomes, the question becomes the relevance of
these to treatment.

Prior work has demonstrated the validity of the SAA in differentiating risk and normative chil-
dren (Crittenden, Kozlowska, & Landini, 2010; Kozlowska et al., 2015; Kozlowska, Scher, &
Williams, 2011; Kozlowska & Williams, 2009; Kwako, Noll, Putnam, & Trickett, 2010). This
study addresses the transition from enacted attachment in the PAA to linguistic representation of
attachment strategy. That is, this study uses the strong evidence of validity of the PAA to establish
the validity of the SAA. By carrying out the PAA when children were 5.5 years of age and then the
SAA 6 months later when the children were 6 years old, we predicted that the SAA would identify
the same attachment strategy as the PAA. However, we also expected that the SAA classifications
would sometimes reflect more detail than the PAA, particularly regarding psychological trauma
and states like depression. In a few cases, life changes were expected to elicit change in strategy in
children. That is, perfect concordance was not expected because of developmental change, change
in the form of assessment, and change in life circumstances.

We expected fairly high proportions of the DMM strategies because epidemiological studies
indicated that 20–30% of the population has a diagnosable disorder at any given time and that low
socioeconomic status (SES) groups are particularly vulnerable (Hagnell, Öjesjö, Otterbeck, &
Rorsman, 1994; Kessler, 1994; Roberts, Attkisson, & Rosenblatt, 1998; Rutter & Rutter, 1993;
Schepank, 1987).

Finally, attachment theory has postulates about the various strategies. These have been tested
for the SSP and AAI and somewhat for the PAA. In this study, we tested the hypothesis that chil-
dren using higher numbered strategies would have been exposed to greater danger, that children
using an A1-2 strategy would inhibit display of negative affect, including anxiety, whereas those
using A3-4, Type C and A/C strategies would express considerable anxiety. Conversely, the inhibi-
tion of Type A individuals would leave them vulnerable to depression, whereas the intense arousal
associated with the Type C strategy would reduce the probability of depression (with Type B show-
ing the least depression).
Hypotheses

Five hypotheses were tested:

1. Children’s 5-year PAA attachment classifications were predicted to match their 6 year SAA classifications clustered as A+ (A3-4), A1-2, B, C1-2, C+ (C3-6) and A/C, thus yielding a $6 \times 6$ comparison.
2. Children drawn from the normative population would more often be classified in low risk classifications (B, A1-2 and C1-2) than children drawn from the clinical group who would more often receive high-risk classifications (A3-4, C3-6 and A/C).
3. Children drawn from the clinical group were expected to more often have indicators of depression and psychological trauma than normative children.
4. Based on epidemiological rates of psychiatric illness, more than a quarter of the low SES normative group were expected to receive risk classifications on the PAA and SAA.
5. Children’s PAA and SAA classifications were expected to be related to measures of stress, child anxiety and maternal and child depression.

Design

The design was a two-group (normative and clinical), two-time (5.5 years and 6 years) multi-construct, multi-method design. There were multiple constructs (e.g. attachment, depression, stress) and these were assessed in multiple ways (self-report, history, interview). The two assessments of attachment were at the upper (PAA) and lower (SAA) age limit of each tool’s range, thus making comparison possible with a minimum of variation due to life changes.

Method

Participants

At the initial contact, the participants were 50 English children between 5.5 and 5.9 years of age and their mothers. Each mother signed an informed consent as approved by the Research Ethics Committee of their Local Authority.

All of the families lived in a small city in an economically depressed area in England. Consequently, even the normative families in this sample were expected to experience more problems than normative families elsewhere. We chose to carry out the research in this city because it was a relatively small and cohesive community that was somewhat isolated from other parts of England (thus reducing movement in and out of the city) and because two of us were well embedded in the community (which facilitated acceptance of the research by the community).

Referrals of normative children were sought from schools where parents were invited to volunteer for the research; referrals of clinically identified children were sought from mental health professionals. The families were mostly working class, with 22 mothers not working, 17 having blue collar jobs, 7 white collar jobs and 4 professional jobs. There was a significant difference between the groups with more working mothers and mothers with white collar jobs in the normative group ($\chi^2(3) = 8.70, p > .034$). A total of 29 were married, with 9 living together, 4 being divorced (and without another partner) and 8 being single. The difference between the groups was significant, with twice as many married mothers in the normative group and almost all single mothers in the clinically referred group ($\chi^2(3) = 10.76, p > .014$). All but one of the children were born healthy and without disability; this one child was in the clinical group. There were 23 girls and 27 boys; the 25 normative children included 12 girls and 13 boys whereas the clinically referred group
consisted of 10 girls and 15 boys (the difference was not significant). There was only one non-White child in the sample. The mothers had an average age of 34 years; there was no group difference.

Of the 50 SAAs, 48 were transcribed and classified; two were lost in transit to the transcriber.

**Schedule of visits**

Each mother–child dyad was seen twice. The Time 1 contact occurred when the children were between 5.5 and 6.0 years old; the children participated with their mothers in a PAA held at a local agency and completed child-report measures of depression and anxiety; the mothers were given a packet of self-report assessments of stress, depression and anxiety to return by mail. The Time 2 visit occurred 6 months later (at 6.0–6.5 years of child age); the children were seen at home and responded to the SAA story cards and child-report measures; their mothers received another packet of measures.

**Assessments**

Except for attachment, most of the assessments were self-report measures. Self-report is known to be vulnerable to social desirability and may also be effected by individuals’ attachment strategies. For example, individuals using a Type A strategy might minimize negative events and emotional states whereas individuals using a Type C strategy might exaggerate these. Only individuals using a Type B strategy would be expected to report fairly accurately on their experience and inner states.

Family stress was assessed using three self-report measures. The Parenting Stress Index (PSI; Abidin, 1995) and Daily Hassles Questionnaire (Department of Health, 2000) are both widely used. They provided a context within which to understand children’s need for protective strategies. More precise data were obtained using the Life Events Inventory (Cochrane & Robertson, 1973) which is a widely used measure of the negative life events; its purpose was to explore the basis in recent family experiences for strategy change or psychological trauma in the children’s SAAs.

Children’s arousal was assessed by the Revised Children’s Manifest Anxiety Scale (Reynolds & Richmond, 1985); it was completed by the child with a researcher at the PAA and SAA meetings.

Maternal depression was assessed with the well-validated Beck Depression Inventory (BDI-II; Beck, Steer & Brown, 1996; Beck, Ward, Mendelson, Mock & Erbaugh, 1961). The purpose was to validate the relevance of maternal low arousal for children’s self-protective strategies.

Child depression was measured using the Child Depression Inventory (CDI; Kovacs, 1992) which was administered in person to children at both Time 1 and Time 2. The CDI requires the reading and comprehension skills of a 6-year-old and based on data from 6- to 17-year-olds (Worchel, Nolan, & Willson, 1987), it is considered applicable to 6-year-olds by Kovacs, even though it is usually used from 7 years onwards. However, being a self-report tool, it is vulnerable to lack of transparency and social desirability.

Attachment was assessed at 5.5 years by the well-validated PAA (Crittenden, 1992; Farnfield, Hautamäki, Norbech, & Sahhar, 2010). The PAA is an enacted assessment that loses its power as children approach 6 years of age and become comfortable being separate from the parents for long periods and able to discuss plans with parents verbally. This diminishing effect of separation to elicit children’s self-protective strategies is greatest among safe children and least pronounced among at-risk children who remain sensitive to the threat of separation longer than other children.

The SAA (Crittenden, 1997–2005) is a relatively new assessment that is being tested for validity in this study. To date, studies demonstrate that it discriminates between normative and clinical
groups (Farnfield et al., 2010; Kozlowska & Elliott, 2014). This study both retested that relation and, more importantly, explored the tie of SAA classifications to the earlier PAA classifications. The important issue is whether the SAA classifications duplicate – and elaborate – the PAA attachment strategies identified for the same children 6 months earlier.

The SAA interviews were audio-recorded and transcribed so as to be annotated for specific discourse markers in each of six memory systems; the discourse markers are derived from the DMM method for analyzing the AAI (Crittenden & Landini, 2011). Using written guidelines (Crittenden, 2009), each SAA protocol is assigned a DMM classification.

**Coding and classification of the PAA and SAA**

There were four coders for the PAAs and five for the SAAs. All had reached reliability on a standardized reliability test and all were blind to all information about the participants and the purpose of the study. In addition, there were periodic checks of coder agreement (1 for each 5 consecutive classifications); these were carried out during the coding process, without coders being aware of which protocols were double coded. Disagreements were resolved through blinded email discussions. Blinding the discussions was intended to reduce the impact of the social desirability of meeting the other person halfway – even if the coder thought that was not the best classification. The kappas for coder agreement were \( \kappa = .56, p < .001 \) on 10 PAAs and \( \kappa = .59, p < .001 \) on 11 SAAs.

**Results**

**Distributions**

Before testing the hypotheses, we examined the distribution of the data on each of our variables. On the PAA, 40% of the protocols were classified as Type A, 20% as Type B, 20% as Type C and 20% as A/C; given that half of the children were clinically referred and all were from a disadvantaged city, the distribution was not unexpected, except possibly with its bias towards A+ as opposed to being evenly distributed between A+ and C+. On the SAA, the distribution was similar (see Table 1).

All 50 mothers reported on negative life events at Time 1, with a mean of 3.26 events and a significant difference between the normative and clinically referred groups \( (F_{(1,48)} = 7.185, p < .01) \). The clinically referred group had more than twice as many negative life events as the normative group (4.48 vs. 2.04). Forty-nine mothers completed the Child Manifest Anxiety Test at Time 1 and 48 at Time 2; there was a nearly significant difference at Time 2 \( (F_{(1,47)} = 3.86, p < .057) \) with the Time 1 means being in the same direction, thus suggesting higher anxiety in the clinically referred group of children.

Only half of the mothers completed the packet of self-report measures, with more having children in the normative group than in the clinically referred group. Length of the measure seemed to reduce the number of respondents, with the PSI having the fewest completions (25) and the BDI the most (35). In every case, the means were higher for the clinically referred group, but the difference was significant only for the Hassle’s intensity and frequency measures \( (F_{(1,26)} = 9.65, p < .005; \) \( F_{(1,26)} = 13.28, p < .001, \) respectively). It is worth noting that having self-report measures completed by the mothers was not an effective way to gather information from this population. Most did, however, comply fully and often quite eagerly when the process included meeting with the researchers.

Forty-four children completed the CDI at Time 1 and 39 at Time 2. There was a significant difference between the normative and clinically referred groups at both time periods: Time 1:
(F(1, 42) = 8.03, p < .007) and Time 2 (F(1, 37) = 6.66, p < .014), with the clinically referred group reporting roughly twice as many signs of depression. The children’s responses at Time 1 and Time 2 were highly correlated (r = .90, p < .000).

**Hypothesis 1**: The hypothesis that each child’s SAA classification would match the PAA classification from 6 months earlier was tested with a chi-square statistic applied to the 6 × 6 set of PAA and SAA classifications. The result was significant (χ^2(25) = 99.291, p > .000). Although there were many empty cells, the overwhelming preponderance of the data, 31 of 47 children, was on the diagonal of matching PAAs and SAAs (κ = .53, p < .000). Considering only major, 4-category (A, B, C, A/C) agreement, there were 34 matches (κ = .58, p < .000) (see Table 1).

**Hypothesis 2**: To test whether clinically referred children would more often show high-risk (A3-4, C3-6 and A/C) strategies than normative children, we ran two simple 2 (referral group) × 2 (low- and high-risk attachment classifications) chi-square tests. One used Time 1 PAA data from all 50 children. Forty children were correctly identified as being normative or clinically referred by the PAA (χ^2(1) = 21.429, p < .000). In addition, however, 10 normative children were given risk PAA classifications; no clinically referred children were given non-risk PAA classifications. The second chi-square used Time 2 SAA data from the 48 children whose SAAs were classified. The results were very similar (χ^2(1) = 21.95, p < .000); nine normative cases had SAA risk classifications; there were no clinically referred children with non-risk attachment classifications.

### Table 1. Crosstabulation of correspondence between PAA and SAA classifications.

<table>
<thead>
<tr>
<th>PAA</th>
<th>SAA</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A+</td>
<td>A1-2</td>
</tr>
<tr>
<td>A+</td>
<td>Count</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>% within PAA</td>
<td>68.8%</td>
</tr>
<tr>
<td></td>
<td>% within SAA</td>
<td>57.9%</td>
</tr>
<tr>
<td>A1-2</td>
<td>Count</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>% within PAA</td>
<td>50.0%</td>
</tr>
<tr>
<td></td>
<td>% within SAA</td>
<td>10.5%</td>
</tr>
<tr>
<td>B</td>
<td>Count</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>% within PAA</td>
<td>.0%</td>
</tr>
<tr>
<td></td>
<td>% within SAA</td>
<td>.0%</td>
</tr>
<tr>
<td>C1-2</td>
<td>Count</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>% within PAA</td>
<td>.0%</td>
</tr>
<tr>
<td></td>
<td>% within SAA</td>
<td>.0%</td>
</tr>
<tr>
<td>C+</td>
<td>Count</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>% within PAA</td>
<td>44.4%</td>
</tr>
<tr>
<td></td>
<td>% within SAA</td>
<td>21.1%</td>
</tr>
<tr>
<td>A/C</td>
<td>Count</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>% within PAA</td>
<td>25.0%</td>
</tr>
<tr>
<td></td>
<td>% within SAA</td>
<td>10.5%</td>
</tr>
<tr>
<td>Total</td>
<td>Count</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>% within PAA</td>
<td>40.4%</td>
</tr>
<tr>
<td></td>
<td>% within SAA</td>
<td>100.0%</td>
</tr>
</tbody>
</table>
Hypothesis 3: Unlike the PAA, the SAA can identify signs of depression and psychological trauma (albeit with less precision regarding the traumatizing event than an AAI). The relation of depression in SAA classifications to referral status was tested in a 3 (depression) × 2 (referral status) chi-square and was highly significant ($\chi^2(2) = 15.68, p < .000$). Of the 23 normative children, two were classified as having partial evidence of depression and one full evidence; of the 24 clinically referred children, 8 showed partial depression and 10 met the full SAA criteria. Psychological trauma followed a similar pattern with a highly significant effect ($\chi^2(1) = 14.727, p < .000$); only 2 of 23 children in the normative group were coded as having psychological trauma whereas 15 of 24 clinically referred children were assigned one or more psychological traumas.

Hypothesis 4: Within the normative subsample, 40% of the children’s PAAs and 42% of their SAAs were classified as A3-4, C3-6 or A/C. Given that A3-4 and C3-4 are less extreme than psychiatric diagnoses and the families came from a disadvantaged population, this seems appropriate.

Hypothesis 5: The hypothesis asked whether there was evidence in self-reports of individual and family functioning that could explain children’s use of risk PAA and SAA strategies and, particularly, depression and psychological trauma (on the SAA). Only the analyses of negative life events and children’s anxiety had complete or nearly complete data. We tested attachment as a 2-group variable (risk versus non-risk classifications) and a 6-group variable (A+, A1-2, B, C1-2, C+ and A/C). Quadratic tests were used with the 6-group attachment classifications to test whether the normative middle categories had lower means than the extremes (i.e. with B as the predicted low point).

Stress: For the PAA, the risk versus non-risk groupings produced no difference on the PSI, but approached significance on the Hassles Intensity measure ($F_{(1, 25)} = 3.59, p < .07$), reached significance on the Hassles frequency measure ($F_{(1, 25)} = 5.48, p < .028$) and approached significance for negative life events ($F_{(1, 48)} = 3.04, p < .09$), with overall higher mean stress for the PAA risk group. The quadratic tests of the 6-way PAA groupings for the PSI, Hassles frequency and Hassles intensity were not significant, but approached significance for negative life events, with Type B having the fewest such events and A+ the most ($F_{(1, 43)} = 3.44, p < .07$).

For the SAA, the risk versus non-risk groupings produced significant or near-significant differences in all the stress-related variables: PSI ($F_{(1, 22)} = 3.19, p < .08$), Hassles intensity ($F_{(1, 24)} = 6.17, p < .02$), Hassles frequency measure ($F_{(1, 24)} = 9.37, p < .005$) and negative life events ($F_{(1, 44)} = 10.11, p < .003$), with overall higher mean stress for the SAA risk group. For the 6-way SAA groupings, the quadratic test for the PSI was not significant whereas the quadratic test for Hassles frequency was statistically significant ($F_{(1, 22)} = 7.63, p < .011$), for Hassles intensity approached significance ($F_{(1, 22)} = 3.30, p < .08$) and for negative life events was significant ($F_{(1, 41)} = 9.67, p < .003$). In each case, Type B had the least stress and A+ and C+ the most.

Anxiety: There was no difference between the PAA risk versus non-risk groups on child-reported anxiety, nor was there a quadratic effect. For the SAA, there were both a risk versus non-risk effect for children’s self-reported anxiety ($F_{(1, 45)} = 7.82, p < .008$) and a 6-group quadratic effect with children assigned to A1-2 reporting the least anxiety and A/C the most ($F_{(1, 42)} = 8.00, p < .007$).

Depression: The 44 children responding on the CDI at Time 1 produced a significant difference when 6-way PAA groupings were used ($F_{(1, 38)} = 5.34, p < .026$) with children using Type B and
C1-2 strategies showing little or no sign of depression and the others showing similarly higher indicators of depression. The 39 children responding on the CDI at Time 2 produced a significant difference when 6-way SAA groupings were used ($F_{(1, 31)} = 8.87, p < .006$); children using Type B and C1-2 strategies showed little or no sign of depression and A+ the highest (Means: $A+ = 13.5; A1-2 = 6.5; B = 3.7; C1-2 = 3.8; C+ = 8.3; A/C = 9.2$).

The results for the 32 mothers’ self-report of depression on the BDI did not approach significance in any grouping of the PAA or SAA.

We also explored the relation of depression and psychological trauma in the SAA with the measures of stress, depression and anxiety. The relation of CDI scores at Time 2 with SAA depression was non-significant. Comparing children assigned full depression to those with partial depression and no depression on negative life events produced a non-significant linearly decreasing set of means. There was no difference in maternal report of negative life events between children with SAA psychological trauma and those without. Because these findings suggested that SAA depression and trauma might be illusory, we compared frequency of these markers by referral group (normative versus clinical). In both cases, the results were highly significant (depression: $X^2(2) = 18.489, p < .000$) and (trauma: $X^2(1) = 14.73, p < .000$) with the clinical group being assigned more depression and trauma. Clearly the SAA and the self-report measures tapped different information, but these findings do not indicate the nature of the difference, possibly because of the low $N$.

Finally, 70% of the SAAs assigned a depressed marker had an A+ strategy with another 15% having an A/C strategy, suggesting the predicted relation of depression to the inhibitory Type A strategy.

**Psychological trauma:** To examine whether the psychological traumas reported by school-aged children were less clear when they reflected parental problems than child problems, we reviewed all the traumas, divided by the referral group. There were 21 different traumas identified by coders, with 4 SAAs from the normative group and 13 SAAs from the clinically referred group. The traumas in the normative group referred to two preoccupying deaths of a sibling at birth, preoccupying physical abuse and denial of physical neglect. The 17 traumas in the clinical group were 15 disorganized, denied, displaced or depressed traumas regarding parents fighting, fathers leaving, the anticipated death of mother, and foster placement and 2 disorganized traumas regarding physical abuse and abandonment of the school-aged child. Strikingly, only two children in the clinical group had neither psychological trauma nor depression. The numbers were too small for statistical comparison.

**Discussion**

**The validity of the SAA**

The central purpose of this study was to test the validity of the SAA from the PAA, given 6 months earlier. Developmentally, this corresponded to the shift from enactment to linguistic representation. Our data strongly supported the short-term longitudinal validity of the SAA classifications as reflecting the same organizations as shown earlier in the PAA.

Our data also replicated previous findings of the validity of the SAA by differentiating the normative and clinical referral groups on (1) child strategy, (2) depression and psychological trauma in the SAAs, (3) negative life events reported by the mothers, (4) less cohesive and financially secure families, (5) greater child anxiety and (6) more child (but not mother)-reported depression
(the group difference held for both 5.5- and 6-year-old children, suggesting the appropriateness of use of the CDI with young children). It is noteworthy that although some normative children used risk attachment strategies, as a group they used few risk strategies and far fewer strategies affected by depression or unresolved trauma. The findings provided a clear pattern of exposure to danger in the clinically referred children’s families; this supports the central tenet of attachment theory (Bowlby, 1969/1983; Crittenden, 1999).

Our study also supported the risk versus non-risk distinction for both the PAA and SAA (with A1-2 and C1-2 considered non-risk classifications). More interesting were the differences between Types A and C. Type A1-2 was similar to B in the apparent lack of anxiety, but similar to A3-4 in terms of signs of depression whereas C1-2 was similar to B in the absence of signs of depression and similar to C3-6 in the presence of anxiety.

**Representation versus enactment of self-protective attachment strategies**

We also demonstrated the superiority of linguistic representation over enacted representation. In every case, the findings for the SAA were stronger and clearer than for the PAA. The move from A/Cs on the PAA to A+ or C+ on the SAA suggests that linguistic representation clarified the enacted representation of the children’s strategy to coders. In addition, three children showed a mismatch across the secure (B) versus anxious (all other classifications) boundary between 5.5 and 6 years of age and a change in how attachment was assessed. Moreover, the SAA added the depth of identifying depressed strategies and strategies interrupted by psychological traumas. Linguistic representation, in fantasy stories and recalled episodes, yielded a breadth of information that enacted strategies cannot.

**Reliability and utility**

Reliability was an issue. The majority of professionals taking the SAA course do not reach reliability and those who did needed extra work beyond the initial course. Partly this is because there is substantial developmental variability between 6 years of age and puberty; most people taking the SAA course were clinicians who had very little developmental training or experience. In addition, most were working clinicians without time for study and practice. Nevertheless, even compared to AAIs (which are known to be difficult to classify), the SAAs are more ambiguous and require that the coder have an ‘ear’ for children’s communication of their meanings. This is especially true for 6-year-olds.

This leads to the issue of utility. It appears that, in the hands of a skilled coder, the SAA yields information that might improve diagnosis and treatment and that is not otherwise available. The issue of psychological trauma is especially interesting. The children in our sample were often troubled by the behaviour of their parents, both towards themselves and towards each other, and of professionals. Because conflict and violence between parents showed in less precise forms of psychological trauma, it appeared to have more generalized impact on children than focused events like physical abuse.

**Constructing good assessments of attachment**

Our review of the development of assessments of attachment points to some crucial criteria for developing valid assessments. Knowing attachment theory intimately undergirds all else; without this, one risks losing construct validity, that is, validity itself. But so does knowing the individuals in the first sample (the sample on which the tool is developed). Without these two bases, one is
flying blind. Once one knows what one wants to observe, a way to elicit it must be devised. Knowing the details of development (mental, emotional and interpersonal) at specific ages is essential for devising the procedure, but so is structuring the procedure to provide comparability across individuals. Furthermore, an assessment of attachment must elicit individuals’ response to threat; if the child does not feel some threat, the self-protective attachment strategy will not be activated. The developers must also be sufficiently able to understand the array of responses to interpret each individual’s meaning. Coding manuals give pointers to help the coder, but in the end, a classification is the coder’s expression of their theory of the assessed person’s state of mind.

Limitations and future directions

The findings of this study are limited by the small sample size and cultural homogeneity of the sample. In particular, the Type A bias in the distribution of the patterns of attachment might reflect an Anglo bias; other cultures, for example Mediterranean cultures, might show a bias to Type C. Missing data reduce the conclusions that can be drawn from those variables, but also point to the limited utility of self-report assessments. Further, the meaning of self-reported information was unclear with several variables showing biases predicted by attachment theory; our study is too small to address these issues adequately. Nor does it fully address which type of assessment (observational with coding or self-report) better describes individuals’ functioning because we did not use self-reported attachment. Given the ease and low cost of self-report measures as compared to the effort and expense of observational assessments of attachment, the latter must have substantially greater validity and utility if they are to be used clinically.

Future work with this sample should examine maternal functioning (we have the mothers’ AAIs and are preparing them for publication) and clinical meaning of the data for individual families (that is, consideration of cases). Studies that compare the validity and utility of different assessments of attachment in a single sample are needed. Finally, with regard to clinical applications, we want to emphasize that assessments of attachment, especially children’s attachment, are not stand-alone tools. To be informative and to protect children and their parents from misapplied treatments, each of the assessments must be viewed in the context of other assessments of that individual and of family members. Attachment is, after all, an interpersonal process.

Acknowledgements

We especially want to thank all of the children and parents who volunteered to help us and who taught us about how much they had to give to others, even when they had struggles of their own. Although we had no funding for this study, we are grateful to the following individuals and agencies for their assistance, without which the study could not have been successful: Margaret Randall, Consultant Child Psychotherapist; Nicky Brewerton, Head, Ramsden Infant School, Barrow; Attachment and Play Therapy; Cumbria Partnership NHS Foundation Trust; Cumbria County Council; Tooby’s Electrical for use of the equipment and IT support; the coders from the Family Relations Institute. Finally, we are grateful for the editorial suggestions from Danya Glaser, Susan Spieker, and Simon Wilkinson.

Funding

This research received no funding from any agency in the public, commercial or not-for-profit sectors.

References


**Author biographies**

Patricia Crittenden studied under Prof. Mary Ainsworth, has been on faculties in North America, Europe, and Australia, and has developed DMM theory and assessments regarding attachment and adaptation across the lifespan. She was given a Career Achievement Award by the European Family Therapy Association in 2004 for ‘Outstanding Contribution to the Field of Child and Family Development.’ She is the Founding Chair of The International Association for the Study of Attachment (IASA).

Katrina Robson is a social worker and child therapist working in England. She has published several papers on DMM applications and is a founding member of IASA. She designed and implemented the DMM-based treatment project, ‘Love Barrow Families’, which is a reorganisation of mainstream services around the needs of multi-problem families.

Alison Tooby is the Lead Child Protection social worker for ‘Love Barrow Families’; she has over 20 years experience of working in a local authority setting and for the NSPCC. She is a founding member of IASA and uses the DMM assessments and knowledge to inform her work with families who face complex difficulties.