International Journal of Behavioral Development
2009, 33 (2), 185–190
http://www.sagepublications.com

ijbd

DOI: 10.1177/0165025409103134

METHODS AND MEASURES

Psychometric properties of the Manchester Child Attachment Story Task: An Italian multicentre study

Lavinia Barone[a], Marco Del Giudice[b], Andrea Fossati[c], Francesca Manaresi[d], Barbara Actis Perinetti[b], Livia Colle[b] and Fabio Veglia[b]

The paper describes a multicentre study of the psychometric properties of the Manchester Child Attachment Story Task in a sample of 230 Italian children aged 4 to 8 years. The task's internal consistency and inter-rater reliability were investigated; in addition, multiple discriminant analysis was used to explore the contribution of individual coding system scale scores to overall categorical attachment classification. The instrument showed acceptable psychometric properties, especially with respect to Disorganization and Coherence scales. However, our results also suggest that some subscales of the coding system could be modified in order to improve reliability. The implications of our results for future research and further test validation are discussed.

Keywords: attachment; childhood; Manchester Child Attachment Story Task; psychometric properties; reliability

In the last two decades, the choice of assessment methods when studying attachment in childhood has become a crucial issue for developmental researchers (see Solomon & George, 1999). Over the years, many new procedures have been developed; currently available methods involve children's responses to pictures depicting separations (e.g., Walsh, Symons, & McGrath, 2004), children's reunion behaviors following separation (Cassidy & Marvin, 1992), and story stem completion tasks (e.g., Bretherton, Prentiss, & Ridgeway, 1990). Whereas researchers can rely on a number of standardized measures – such as the well-known Strange Situation Procedure – when studying infants and toddlers (12 to about 20 months), attachment theory is less specific regarding suitable measures of security after the second year of life and beyond. In childhood, the smaller number of situations perceived as threatening (and thus appropriate for eliciting attachment), the broader and multidimensional concept of parents' accessibility (extending beyond actual proximity and contact seeking, which are typical of infancy), and the child's achievements in the behavioral, symbolic, cognitive, and linguistic domains, all contribute to make attachment assessment a challenging but crucial task (for reviews and discussion see Kerns, Schlegelmilch, Morgan, & Abraham, 2005; Kerns, Tomich, Aspelmeier, & Contreras, 2000). Despite the above-mentioned shortcomings, story stem techniques have become a widely accepted method for indirectly assessing children's internal working models of attachment relationships (Robinson, 2007). These methods – such as the MacArthur Story Stem Battery -MSSB (Bretherthon, Oppenheim, Emde, and the MacArthur Narrative Working Group, 2003), the Attachment Story Completion Task – ASCT (Bretherton, Ridgeway, & Cassidy, 1990) and the Manchester Child Attachment Story Task – MCAST (Green, Stanley, Smith, & Goldwyn, 2000) – use doll figures to elicit children's narratives in response to attachment-related themes. Children's attachment representations are then inferred from their verbal and nonverbal responses, by applying specific coding systems to several dimensions of behavior and narrative content and style. The coding systems of the ASCT and MCAST consist of a set of scales measuring various dimensions of narrative attachment representations; scores on the coding scales are then combined to obtain a 4-way classification of attachment patterns (i.e., A: insecure avoidant; B: secure; C: insecure ambivalent; and D: disorganized). Clearly, the reliability of such coding systems is central to establishing doll-play techniques as useful instruments for assessing attachment representations in childhood.

The aim of this paper is to investigate and discuss some basic psychometric features of a relatively new story stem technique – the Manchester Child Attachment Story Task (Green et al., 2000) – which combines criteria derived from both infant and adult research (i.e., Strange Situation Procedure and Adult Attachment Interview) and is deemed suitable for children aged from 4 to 8 years. The instrument we investigated, mainly derived from the Attachment Story Completion Task of Bretherton and colleagues (1990), benefits from a broadly articulated coding system which, in principle, could allow the reliable measurement of many specific dimensions pertaining to the challenging theoretical construct of attachment in childhood (e.g., mentalization or specific facets of disorganization such as controlling behavior). Although this technique

[a]Pavia University, Italy. [b]Turin University, Italy. [c]Vita-Salute University, Italy. [d]ARPAS, Italy.

Correspondence should be sent to Lavinia Barone, Department of Psychology, Pavia University, Piazza Botta 6, Pavia 27100, Italy; e-mail: lavinia.barone@unipv.it

represents a promising assessment method (Goldwyn, Stanley, Smith, & Green, 2000), no large sample study to assess its psychometric properties has yet been performed.

Since the Manchester Child Attachment Story Task is extensively used in Italy and other European countries in both clinical and research settings, we felt the need to carry out a large-sample multicentre study (comprising four Italian research groups) aimed at investigating the basic psychometric properties of its coding system. In particular, our study was aimed at investigating:

1 the inter-coder reliability on 4-way overall attachment classifications;
2 the agreement between 4-way coding of individual vignettes and the overall attachment classification, in order to evaluate the relative contribution of each vignette to the final test classification;
3 the consistency of the scores on each coding system scale across the four attachment vignettes, and the internal consistency (i.e., reliability) of the scores on each scale averaged over the four attachment-related vignettes;
4 the ability of the coding system individual scales to discriminate the four attachment categories that are generated by each vignette (i.e., ABCD). In a sense, this is akin to item analysis in classical test development.

Method

Participants

Two-hundred and thirty low-risk children (50% males) between the ages of 53 and 96 months (M = 80.8 months, SD = 13.9) participated in the study. Participants were recruited from local preschools (33%) and primary schools (67%) of four Italian metropolitan areas and were tested at school in an individual and quiet setting. The children were primarily from two-parent homes (94% married) and none reported any major physical or behavioral problems. Data on socioeconomic status, collected using the Hollingshead scale (1975), were available for all but one group (SES mean = 61.83; SD = 19.93). One of the samples consisted of N = 60 children randomly selected from the larger sample described in Del Giudice (2008).

Given that neither gender ($\chi^2(3) = 4.78$, $p > .15$), sample size ($\chi^2(3) = 1.44$, $p > .60$), nor socio-cultural differences were significant among the four sites, all statistical analyses were performed on the combined sample.

Procedure and measures

Each child is introduced to the play materials (including a dolls house with furniture) and asked to choose dolls to represent her- or himself and a primary caregiver (chosen by the interviewer). A warm-up "breakfast" vignette is proposed to each child, both as an introduction to the procedure and in order to provide a non-attachment baseline for the child's behavior. Then four attachment-related vignettes are presented. The first one ("nightmare") presents the child with a situation where the child-doll wakes up at night, all alone, after having a nightmare; in the second ("hurt knee"), the child-doll is playing outside, falls over and cuts his/her knee which starts bleeding; in the third ("tummy ache") the child-doll is watching TV and suddenly starts having a tummy ache. Finally, in the fourth vignette ("shopping"), the child-doll gets lost while shopping with her/his mother-doll in a large mall. A final, free-play vignette is aimed at cooling down the emotional arousal elicited by the attachment-related vignettes. Only the four attachment-related vignettes are scored and used for attachment classification. The stories' coding scales fall into four broad dimensions (Green, Stanley, Goldwyn, & Smith, 2000–05):

1 attachment-related behaviors (e.g., proximity seeking and assuagement strategies);
2 narrative coherence, i.e., the four coherence principles of Grice's maxims of discourse (1975);
3 disorganized phenomena (including five disorganization facets: chaos; no identifiable strategy; use of multiple and incompatible strategies; episodic disorganization, i.e., narrative disruption; control of caregivers either solicitous or coercive; and finally presence of bizarre themes without resolution).
4 mentalization skills (i.e., the child's awareness of the states of mind of characters in the story and in his/her metacognition).

These four thematic dimensions are scored through 21 rating scales (all ranging from 1 to 9, with the exception of the scales labeled "engagement", "arousal", "turn-taking", and "mentalization", which range from 0 to 3). An overall "strategy of assuagement" is then obtained by comparing the child's scale scores with prototypical profiles, and each vignette is categorized in the standard 4-way classification: A (avoidant); B (secure); C (ambivalent); and D (disorganized). The four vignettes' categories are then combined to form the child's overall attachment classification, according to a majority rule: if two or more vignettes receive the same attachment classification, this becomes the child's overall attachment category.

Protocols were scored by first, second, fourth, and fifth author respectively, all trained in the instrument coding system and checked for reliability by Dr. Jonathan Green (University of Manchester). Participants' attachment classifications were the following: 63% secure (B), 16% insecure-avoidant (A), 10% insecure-ambivalent (C), and 11% insecure-disorganized (D). Both age (ANOVA: $F(3,226) = 8.86$, $p < .001$) and gender ($\chi^2(3) = 14.26$, $p < .005$) were not homogeneously distributed across the four attachment patterns. In particular, Bonferroni contrasts showed that avoidant children (A) were significantly younger than secure (B) or ambivalent (C) children, whereas both A and C were significantly older than disorganized (D) subjects. Furthermore, children categorized as ambivalent (C) were more frequently females (76%), whereas A and D children were more often males (69% and 63%, respectively). See Del Giudice (2008) for a more detailed discussion of this finding.

Analytic plan

In order to answer our first research question, Cohen κ coefficient was computed for assessing inter-rater agreement reliability of the 4-way overall attachment categories (indeed, Cohen κ is analogous to a one-way ANOVA intraclass correlation coefficient when data are discrete; Fleiss, 1973). Our second question was investigated by measuring the chance-corrected agreement between 4-way coding of individual vignettes (i.e., ABCD) and the overall attachment classification

(i.e., ABCD). This was also done by computing Cohen κ coefficients.

Since the Manchester Child Attachment Story Task generates absolute evaluations (i.e., the attribution of a given subject to the A, B, C, or D classification does not require the comparison of his/her scores with a normative distribution), we sought to answer our third research question (concerning consistency of the scores of each coding system scale across the four attachment vignettes) by computing intraclass correlation coefficients based on a one-way random effect ANOVA (Shrout & Fleiss, 1979).

Finally, linear discriminant analysis was used to answer our last question, concerning how individual scales predict the attachment classification in the individual vignettes and in the overall categorization. Discriminant analysis involves obtaining a linear combination of predictors (coding system scales) that best discriminate among groups (ABCD classifications); in our case, the number of possible linear combination of predictors (discriminant axes or functions) along which the attachment group could vary was three. In order to evaluate the significance of the contribution of each scale to attachment classification, the best predictors were selected using a stepwise algorithm based on Wilks Λ minimization (i.e., the lower the value of Wilks Λ, the greater the difference between the attachment groups), with significance level set to $p > .05$. This analysis is similar to classical item-total analyses in that they aim to assess whether in each vignette all instrument scales contribute equally and significantly to the overall attachment classification, or if there are some of them which appear ineffective or redundant.

Results

Inter-rater agreement and chance-corrected agreement between individual vignettes and the overall attachment classification

Inter-rater reliability was tested on 50 overall classifications (21.7% of the sample). Raw agreement on security vs. insecurity was 86% (Cohen's $\kappa = .72$, $p < .001$), and 78% on the 4-way classification (A, B, C, D). Cohen's κ values showed that, as a whole, the 4-way attachment strategies identified in the individual test vignettes were moderately associated with the overall test classification, with some degree of variation between performance for vignette 1 (nightmare) and the remaining three vignettes (see Table 1). Interestingly, only moderate agreement was observed among the predominant strategies of the individual vignettes. The predominant strategies observed in vignette 1 were the least replicated across the remaining vignettes.

Table 1
Chance-corrected association between individual vignettes and overall attachment classification (Cohen κ)

	Overall	*Vignette 1*	*Vignette 2*	*Vignette 3*
Vignette 1 (nightmare)	.55	1.00		
Vignette 2 (hurt knee)	.64	.47	1.00	
Vignette 3 (tummy ache)	.69	.39	.57	1.00
Vignette 4 (shopping)	.67	.40	.44	.54

Score consistencies of the coding system scales across the four vignettes. Four scales showed poor consistency of the individual scores across vignettes, as assessed by intraclass correlation (ICC; Table 2): Proximity-seeking, Self-care, Reversal, and Exploratory play. The remaining 17 scales showed intraclass correlations between .51 and .70, suggesting moderate, albeit respectable consistency values. Moreover, these values increased in magnitude when considering score reliability averaged across the four vignettes; there, ICC coefficients were highly satisfactory for all the scales (ranging from .71 to .92), with the only exception of the Self-care scale.

Coding system scales discriminant validity with respect to attachment classification. In order to assess the contribution of each scale to 4-way attachment classification, we ran a stepwise discriminant analysis as described earlier. All the final models generated by the stepwise algorithm were highly significant: vignette 1, Wilks $\Lambda = 0.30$, $p > .001$, % of correctly classified subjects = 76.8; vignette 2, Wilks $\Lambda = 0.20$, $p > .001$, % of correctly classified subjects = 82.1; vignette 3, Wilks $\Lambda = 0.17$, $p > .001$, % of correctly classified subjects = 81.7; vignette 4, Wilks $\Lambda = 0.26$, $p > .001$, % of correctly classified subjects = 79.8. Jackknifed classifications did not differ appreciably from these estimates.

According to the stepwise algorithm, only sub-sets of scales significantly (i.e., $p > .05$) minimized Wilks Λ statistics, and they could be linearly combined to predict the overall attachment classification. The structure of the three discriminant axes – i.e., the three linear combination of the coding system scales in each vignette that significantly discriminates among 4-way attachment classifications – is shown in Table 3. Structure coefficients are correlations between the individual coding system scales and the discriminant axes; they are indexes to help define which scales are most important in discriminating among attachment 4-way classifications.

Table 2
MCAST scales: Intraclass correlations for individual and average scores across the four vignettes

MCAST scales	*Individual score ICC*	*Average score ICC*
1. Engagement	.70	.90
2. Arousal	.57	.84
3. Turntaking	.73	.91
4. Proximity-seeking	.44	.76
5. Self-care	.23	.55
6. Displacement	.61	.86
7. Role-reversal	.41	.74
8. Conflicted behavior	.51	.80
9. Carer sensitivity	.60	.86
10. Carer intrusiveness	.58	.85
11. Assuagement (child)	.50	.80
12. Assuagement (observer)	.57	.84
13. Exploratory play	.38	.71
14. Play content	.64	.88
15. Quality	.72	.91
16. Quantity	.70	.90
17. Relevance	.70	.90
18. Manner	.73	.91
19. Overall coherence	.74	.92
20. Mentalization	.52	.81
21. Episodic disorganization	.62	.87

Table 3
Discriminant validity of MCAST scales to attachment classifications: discriminant function structure coefficients

	Discriminant axes (Linear combinations of predictors)		
MCAST scales	*1*	*2*	*3*
Vignette 1 (nightmare)			
1. Turntaking	.44	.03	.17
2. Displacement	.48	.24	–.26
3. Role-reversal	–.41	.10	–.12
4. Conflicted behavior	–.52	.24	–.15
5. Carer sensitivity	–.10	–.76	.69
6. Carer intrusiveness	–.09	.62	–.03
7. Assuagement (child)	.39	–.20	–.21
8. Play content	.43	.42	.82
Vignette 2 (hurt knee)			
1. Engagement	–.12	.16	–.41
2. Proximity-seeking	.14	.36	.05
3. Conflicted behavior	–.17	.63	–.00
4. Assuagement (child)	.25	–.68	–.11
5. Exploratory play	–.17	.08	.41
6. Play content	–.23	–.24	.76
7. Quality	.04	–.10	.93
8. Manner	.54	.06	.28
9. Episodic disorganization	–.26	–.27	.46
Vignette 3 (tummy ache)			
1. Engagement	.22	.14	–.34
2. Conflicted behavior	.11	.43	–.39
3. Carer intrusiveness	–.16	.39	.30
4. Assuagement (child)	–.11	–.62	.12
5. Play content	.40	.32	.53
6. Quality	.13	.00	.88
7. Relevance	–.55	.31	.20
8. Episodic disorganization	.46	–.22	.69
Vignette 4 (shopping)			
1. Role-reversal	–.09	.36	.34
2. Conflicted behavior	–.19	.30	.66
3. Carer sensitivity	.34	.42	.05
4. Assuagement (observer)	–.01	.50	–.43
5. Play content	–.32	.30	.07
6. Manner	.53	.44	.09
7. Episodic disorganization	–.05	.36	–.67

Roughly fifty percent of the coding scales were non-redundant in significantly predicting the final attachment classification. With the exception of Turn-taking, Arousal, Displacement, Proximity-seeking, Self-care, Assuagement (observer), Exploratory play, and Mentalization, all of the other scales were present as predictors in two or more vignettes. Interestingly, Conflicted behavior significantly predicted the final classification in all vignettes, and Episodic disorganization and Assuagement (child) scales were significant predictors of the 4-way attachment classification in three out of four vignettes. Among coherence subscales, quality and manner were the best predictors across vignettes.

Discussion

In this study we carried out a large sample investigation of the psychometric properties of the Manchester Child Attachment Story Task, concerning in particular its internal consistency and inter-rater reliability. While the present study was not designed to validate the test, internal consistency and reliability are essential conditions for validity; the study can thus be seen as a preliminary effort toward the large-scale validation of the instrument coding system.

The results showed that the reliability measured on individual scales was moderate. However, the fact that the average scores across the vignettes were generalized for all (but one) coding scale supports the instrument's overall reliability. Discriminant analysis also indicated that the coding system scales (collectively taken) have good predictive value with respect to 4-way attachment classification. While the overall predictive performance of the scales is satisfactory, not all of them appear to contribute equally to the prediction. Some of them, such as Assuagement (child), Carer sensitivity, Episodic disorganization and Conflicted behavior were selected as significant predictors in most vignettes; on the other hand, some only predicted classification in one vignette (e.g., Exploratory play, Displacement), and the Self-care and Mentalization coding scales never proved to be significant predictors. The odd behavior of these scales is also apparent from the low ICC consistency scores of one of them (Self-care), which may suggest problems with the manual scoring instructions. Moreover, range restriction in some of them (e.g., Turn-taking, Arousal, and Mentalization) is likely to have played a role in their poor predictive power either because of narrow coding range or because of low frequency of target responses. We suggest that one reason for these findings could be the design of the coding system manual (Green et al., 2000–05). The categorical attachment classification of each vignette is drafted according to a prototype-based system, which does not include detailed cutoff points for the individual coding scales. This could explain the relatively poor contribution of some scales, which could be improved by linking them to the final classification in a stricter manner (e.g., by empirically deriving detailed cutoff scores in addition to prototypes).

We found satisfactory consistency between individual vignettes and overall 4-way classification. Of the four vignettes, vignette 1 (nightmare) showed the least association with overall classification, mainly because of some disorganized subjects being misclassified as secure. This might be due to some specific aspect of the vignette's content, and/or to the vignette's initial position in the test.

Our analysis of the inter-coder agreement statistics shows that 4-way classifications have good reliability (about 80% inter-coder agreement). This result is consistent with the only psychometric study of this instrument currently available (Goldwyn et al., 2000).

One of the most interesting features of the Manchester Child Attachment Story Task is its fine-grained assessment of disorganization severity via the episodic Disorganization scale. The instrument sensitivity in this respect has been already shown in the original validation study (Goldwyn et al., 2000). In our analysis, this scale was a reliable predictor of 4-way classification in three out of four vignettes (all but vignette 1), which further supports its validity. In the current version of the manual, coherence is treated as a global score and does

not enter directly in the overall 4-way classification, rather serving as a "subsidiary" narrative quality scale. Our results, however, suggest that individual Coherence scales (i.e., Quantity, Quality, Relevance, and Manner) and/or their combination might prove useful in identifying specific attachment patterns.

Limitations and future directions

Our study addressed an age range (early and middle childhood) in which there have been few large-scale, rigorous inquiries about the reliability of attachment assessment methods. Most seminal studies have relied on relatively small samples, usually around 30–100 children (e.g., Ammaniti, van Ijzendoorn, Speranza, & Tambelli, 2000; Granot & Mayseless, 2001; Green et al., 2000; see however the larger validation studies by Finnegan, Hodges, & Perry, 1996 and Target, Fonagy & Shmueli-Goetz, 2003). Our multicentre design – comprising four low-risk samples from different research groups in Italy – provided a large, moderately heterogeneous sample of preschool and primary school children, thus offering a good benchmark for testing the instrument's psychometric properties. Our findings showed that the Manchester Child Attachment Story Task is a promising instrument with acceptable internal consistency and reliability; however, they also suggest that its coding system could be further improved to increase some reliability coefficients and to reduce redundancy among scales.

Clearly, our study also has a number of limitations. First of all, we did not address the issue of concurrent and external validity. As stated above, reliability is a necessary but not sufficient condition for validity. Future studies need to correct this limitation by adding external criteria in order to fully validate the instrument. Second, our exclusive focus on basic psychometric properties prevented us from addressing important theoretical issues concerning the meaning and measurement of attachment in childhood (see Solomon & George (1999) and Kerns et al. (2005) for in-depth discussion of these topics).

Our psychometric analysis suggests that some improvements to the coding system may be possible. In particular, some scales appear to be redundant and/or less reliable than others (see, for example, the Mentalization and Self-care scales). Collecting and pooling data from larger samples will allow for more detailed analysis, suggesting which scales could be removed or reworked. Another major improvement would be calculating cutoff scores for the rating scales, to be used in addition to prototype-matching. In addition, we suggest that it would be informative to randomize the presentation order of the four attachment-related vignettes, at least for research purposes. Given the differences in predictive power between vignettes that we found in the present study, this would allow for a more rigorous assessment of the homogeneity of the attachment-related situations proposed in the task. Special attention should also be given to the role played by Coherence scales. Coherence has proved to be a central construct both in adult interviews (e.g., Beijersbergen, Bakermans-Kranenburg, & van Ijzendoorn, 2006) and in interviews for older children and adolescents (see Ammaniti et al., 2000; Target et al., 2003). Our data suggest that this dimension may play a similarly important role in early childhood as well, and may be applied with success to doll-play assessment methods. Future revisions of the coding manual could include Coherence "profiles" (e.g., low quality and relevance, coupled with high quantity and manner) specifically linked to attachment classification.

We believe that, at the present state of research, there is still much work to do in the area of attachment assessment in childhood. In particular, few instruments have been extensively tested for reliability and consistency, as we have tried to do with the Manchester Child Attachment Story Task in the present study. However, complex coding systems usually need to undergo rigorous psychometric analysis before they can be successfully applied to research or clinical settings. Large-scale analyses of this kind (possibly involving multicentre designs and multiple raters) can have a crucial role in suggesting revisions and improvements, thus fostering the continued development of attachment tests and increasing their scientific usefulness.

References

Ammaniti, M., van Ijzendoorn, M.H., Speranza, A.M., & Tambelli, R. (2000). Internal working models of attachment during late childhood and early adolescence: an exploration of stability and change. *Attachment & Human Development*, *2*, 328–346.

Beijersbergen, M.D., Bakermans-Kranenburg, M.J., & van Ijzendoorn, M.H. (2006). The concept of coherence in attachment interviews: Comparing attachment experts, linguistics, and non-experts. *Attachment & Human Development*, *8*(4), 353–369.

Bretherton, I., Oppenheim, D., Emde, R.N., & the MacArthur Narrative Working Group (2003). In R.N. Emde, D.P. Wolf, & D. Oppenheim (Eds.), *Revealing the inner worlds of young children: The MacArthur Story Stem Battery and parent–child narratives* (pp. 381–396). New York: Oxford University Press.

Bretherton, I., Prentiss, C., & Ridgeway, D. (1990). Family relationships as represented in a story-completion task at thirty-seven and fifty-four months of age. *New Directions for Child Development*, *48*, 85–105.

Bretherton, I., Ridgeway, D., & Cassidy, J. (1990). Assessing internal working models of the attachment relationship: An attachment story completion task for 3-year-olds. In M.T. Greenbergh, D. Cicchetti, & E.M. Cummings (Eds.), *Attachment in the preschool years* (pp. 273–308). Chicago, IL: University of Chicago Press.

Cassidy, J., & Marvin, R.S. (1992). *Attachment organization in three and four year olds: Procedures and coding manual*. University of Virginia: unpublished coding manual.

Del Giudice, M. (2008). Sex-biased ratio of avoidant/ambivalent attachment in middle childhood. *British Journal of Developmental Psychology*, *26*, 369–379.

Finnegan, R.A., Hodges, E., & Perry, D. (1996). Preoccupied and avoidant coping during middle childhood. *Child Development*, *67*, 1318–1328.

Fleiss, J.L. (1973). *Statistical methods for rates and proportions*. Oxford: John Wiley & Sons.

Goldwyn, R., Stanley C., Smith V., & Green, J. (2000). The Manchester Child Attachment Story Task: relationship with parental AAI, SAT and child behaviour. *Attachment & Human Development*, *2*(1), 71–84.

Granot, D., & Mayseless, O. (2001). Attachment security and adjustment to school in middle childhood. *International Journal of Behavioral Development*, *25*, 530–541.

Green, J.M., Stanley, C., Goldwyn, R., & Smith, V. (2000–05). *Coding manual for the Manchester Child Attachment Story Task*. Manchester University: unpublished manual.

Green, J.M., Stanley, C., Smith, V., & Goldwyn, R. (2000). A new method of evaluating attachment representation in young school-age children – the Manchester Child Attachment Story Task. *Attachment & Human Development*, *2*(1), 42–64.

Grice P. (1975). *Logic and conversation*. In P. Cole & J.L. Morgan (Eds.), *Syntax and semantics. Vol. 3. Speech acts*. New York: Academic Press.

Hollingshead, A. (1975). *Four factors index of social position*. Yale University: unpublished manuscript.

Kerns, K.A., Schlegelmilch, A., Morgan, T.A., & Abraham, M.M. (2005). Assessing attachment in middle childhood. In K.A. Kerns & R.A. Richardson (Eds.), *Attachment in middle childhood* (pp. 46–70). New York: The Guilford Press.

Kerns, K.A., Tomich, P., Aspelmeier, J., & Contreras, J. (2000). Attachment-based assessments of parent–child relationships in middle childhood. *Developmental Psychology*, *36*, 614–626.

Robinson, J.L. (2007). Story stem narratives with young children: Moving to clinical research and practice. *Attachment & Human Development*, *9*, 179–185.

Shrout, P.E., & Fleiss, J.L. (1979). Interclass correlations: Uses in assessing reliablity. *Psychological Bulletin*, *86*, 420–428.

Solomon, J., & George, C. (1999). The measurement of attachment security in infancy and childhood. In J. Cassidy & P.A. Shaver (Eds.), *Handbook of attachment: Theory, research and clinical applications* (pp. 287–318). New York: The Guilford Press.

Target, M., Fonagy, P., & Shmueli-Goetz, Y. (2003). Attachment representations in school-age children: the development of the child attachment interview (CAI). *Journal of Child Psychotherapy*, *29*, 171–186.

Walsh, T.M., Symons, D.K., & McGrath, P.J. (2004). Relations between young children's responses to the depiction of separation and pain experiences. *Attachment & Human Development*, *6*, 53–71.