Measuring Attachment Security: Concurrent and Predictive Validity of the Parental Attachment Q-set

MARION VAN DAM
MARINUS H. VAN IJZENDOORN
Department of Education
University of Leiden, The Netherlands

ABSTRACT. In this study, concurrent and predictive validity of Waters and Deane's (1985) Attachment Q-set was tested. We hypothesized that outcomes of the Attachment Q-set would converge with Strange Situation behavior and classification and that maternal responsiveness would correlate with security (as measured through the Attachment Q-set) but temperamental characteristics would not have a direct relation to security. Thirty-nine children ($M = 18$ months) and their mothers participated in the Strange Situation procedure and in a free play session in which maternal responsiveness was measured. Mothers completed the Attachment Q-set twice and also filled out Bates' Infant Characteristics Questionnaire (Kohnstamm, 1984). Results indicated a lack of concurrent validity: Strange Situation and Attachment Q-set outcomes did not converge. Although responsiveness was related to attachment as measured through the Strange Situation, it showed no overlap with security as measured through the Attachment Q-set. Security also appeared to be related in an unexpected way to difficultness and adaptability as measured through the ICQ. Therefore, we were unable to establish the predictive validity of the Attachment Q-set.

UNTIL SOME YEARS AGO, researchers had to apply the well-known Strange Situation procedure to measure attachment security (Ainsworth, Blehar, Waters, & Wall, 1978). Recently, however, Waters and Deane (1985) introduced the Attachment Q-set as a viable alternative. This new measure had the following advantages: It can be applied in a broader age range (1 to 3 years old), it can be scored much more efficiently and reliably than the Strange Situation; its ecological validity is greater because it is applied in a natural setting; it takes discriminant validity into account by measuring not only attachment security but also such constructs as dependency and sociability; and it contains a measure of observers' tendencies to record socially
desirable characteristics in children’s behaviors. Although the Attachment Q-set and its scoring criteria have been published in some detail (Waters & Deane, 1985), no information about its concurrent validity with the Strange Situation has been made available. The Attachment Q-set should only be considered a viable alternative to the classical Strange Situation procedure if the results of both instruments show some convergence.

A new measure for attachment security should also show some predictive validity, that is, predictable relations to external criteria or variables (Nunnally, 1978). One of the core elements of attachment theory implies a relation between the caregiver’s responsiveness and a child’s attachment security (Van IJzendoorn & Tavecchio, 1987). From attachment theory the prediction may be derived that caregivers who respond more promptly and adequately to the child’s signals (Ainsworth, Bell, & Stayton, 1974) more often build a secure relationship with their child (Ainsworth et al., 1978). Although not all studies on the relation between attachment and responsiveness have reported large effect sizes (see Goldsmith & Alansky, in press, and Lamb, Thompson, Gardner, & Charnov, 1985, for an extensive overview), the data obtained do tend to confirm the expected relation between Strange Situation outcome and responsiveness. Responsiveness may, therefore, be considered an important criterion for determining the predictive validity of the Attachment Q-set.

In attachment theory, it has been hypothesized that temperamental differences—differences in relatively enduring, cross-situationally stable aspects of behavioral style (Kohnstamm, 1986)—cannot tip the balance in classifying infants’ behavior in the Strange Situation into one of the main categories of attachment security (Bretherton, 1985; Sroufe, 1985). Strange Situation classifications should not be contaminated by temperamental differences between children, though certain groupings of children’s subclassifications (e.g., A1, A2, B1, B2 versus B3, B4, C1, C2) may be explained in part by temperament (Belsky & Rovine, 1987). In studies using a parent-report measure to assess temperament, relations between temperament and attachment classifications have been very small or nonexistent (e.g., Bates, Maslin & Frankel, 1985; Egeland & Farber, 1984; Thompson & Lamb, 1982). Although data from different studies do not always seem consistent (see Goldsmith & Alansky, in press; Goldsmith, Bradshaw, & Rieser-Danner, 1986; and Lamb et al., 1985, for an extensive review), we consider absence of shared variance between attachment and temperament as indicating (discriminant) validity.

We gratefully acknowledge the assistance of Corinne van der Heuvel, Hans Plomp, Annemarth van der Vossen, and Hylida Woudstra in collecting and coding the data. We thank Everett Waters, who provided us with the Attachment Q-set.

Requests for reprints should be sent to Marinus H. van IJzendoorn, Department of Education, University of Leiden, P.O. Box 9507, NL-2300 RA Leiden, The Netherlands.
In short, we tested three hypotheses regarding the validity of the Attachment Q-set. First, we hypothesized that the Q-set measure should be related to the outcomes of the classical Strange Situation procedure. Second, we expected responsiveness to correlate with attachment security as measured through the Strange Situation and through the Attachment Q-set. Third, we hypothesized that temperamental differences would not explain much variance in attachment security as measured by the two instruments.

Method

Subjects

Thirty-nine families participated. Subjects were recruited through birth announcements in a local paper. The mean age of the infants was 18 months (± 2 weeks); 19 were boys, 20 were girls. The socioeconomic status of the families participating \((M = 4.2; SD = 1.4)\) was determined according to the Van Westerlaak, Kropman, and Collaris (1975) occupational index, in which occupational level is scored on a scale extending from unskilled labor (1) to high-level and academic occupations (6).

Instruments and Procedure

Attachment Q-set. We used the simple parental version of the Attachment Q-set constructed by Waters and Deane (1985), containing 75 items (Deane, 1985). These items must be sorted into nine piles whose designations range from most characteristic to least characteristic of a particular child. This is accomplished in three steps, leading to a final sort, which conforms to a symmetrical, unimodal distribution (for details, see Waters & Deane). Mean correlations between observers' and parent's versions of the Attachment Q-set were high (.80; Waters & Deane); differences were "clearly examples of the mother having better access to the behavior than the observers did" (p. 59). The original English version was independently translated by three translators, consensus about the correct translation was arrived at by discussions, and the result was translated back into English. Divergences between this translation and the original version were minimized by polishing the translation into Dutch. In addition, some items with double negatives were changed to avoid confusing the parents too much (items 6, 8, 13, 16, 22, 25, 39, 40, 41, 44, 49, 54, 59, 62, 65 and 74).

We asked mothers to sort the Q-set cards two times, before and after visiting the laboratory to participate in the Strange Situation; they were helped by research assistants, who explained the procedure and the behavioral descriptions in response to questions from the mothers.

Q-set outcomes were correlated with criterion sorts for security, dependency, sociability, and desirability (Waters & Deane, 1985), after recoding
the items with double negatives. The criterion sorts were for 12-months-olds. The correlations served as scores on the variables security, dependency, sociability, and desirability. A high security score indicates that a child was rated as more anxious; a high dependency score indicates that a child was more independent; a high sociability score means that a child was less sociable; and a high desirability score means that parents gave fewer socially desirable answers. Test-retest reliability scores for security, dependency, sociability, and desirability (over a period of about 10 days) were .75, .86, .78, and .82, respectively. The mean score on security was −.405 (SD = .136), on dependency, .141, (SD = .241), on sociability, −.327 (SD = .184); and on desirability −.410 (SD = .174).

**Infant Characteristics Questionnaire (ICQ).** The Dutch version of Bates' ICQ was used to assess temperament (Bates et al., 1985). Kohnstamm (1984) revised the scale for use with Dutch subjects and studied validity and reliability issues in a large sample of 7,000 Dutch parents. In his study, as well as in the study of Tavecchio and Van IJzendoorn (1987) on 413 families, the following dimensional structure was found:

1. **Difficultness**, which indicates how often a child cries or fusses, how irritable he or she is, and how much attention he or she generally demands from caregivers. This dimension also reflects the mother's opinion of how difficult the child is, for other parents as well as for herself. A low score indicates a relatively easy temperament. (In this study: $M = 21.2, SD = 4.2; \alpha = .69$.)

2. **Adaptability** reflects the child's ability to adjust him- or herself to new circumstances, for example, new environments or strangers. A low score indicates a high degree of adaptability. (In this study: $M = 10.7, SD = 4.1, \alpha = .87$.)

3. **Cuddly** indicates whether or not the child likes to be picked up or cuddled and how often he or she insists on being cuddled by nestling close to the caregiver. A low score indicates that the child dislikes being cuddled. (In this study: $M = 8.9, SD = 2.4, \alpha = .66$.)

4. **Persistence** reflects the child's obedience and docility. A high score indicates a relatively disobedient, "naughty" type of behavior. (In this study: $M = 8.2, SD = 2.5, \alpha = .76$.)

5. **Mood** indicates how often the child is cheerful and what may be considered his or her typical mood (cheerful or sober). High scores point to a dejected mood ($M = 4.0; SD = 1.5$). Alpha reliability was .86 (see Kohnstamm, 1984, and Tavecchio & Van IJzendoorn, 1987, for details).

In this study, we used the same five dimensions as temperament variables. Correlations between scales were low ($M$ correlation = .21), and alpha reliabilities (see above) were sufficient (Nunnally, 1978).
The Strange Situation. The Strange Situation procedure for measuring the quality of mother–infant attachment consists of eight episodes, the last seven of which should ideally last 3 min each. The infant is confronted with a strange environment (i.e., the playroom in the laboratory), with a strange person, and with two separations from his or her caregiver (for details, see Ainsworth et al., 1978). The child’s behavior is scored on six 7-point rating scales: proximity and contact seeking, maintenance of contact, resistance, avoidance, search behavior, and distance interaction. After examining the pattern of scores in the two reunion episodes, observers classify mother–infant dyads into one of the following attachment categories: anxiously avoidant (A), secure (B), or anxiously resistant (C). These three main groups may be further divided into eight subgroups (A1, A2, B1, B2, B3, B4, C1, and C2). In our sample, 14 dyads were classified as A, 21 dyads were classified as B, and 4 dyads were classified as C (Van IJzendoorn & Kroonenberg, 1988).

Two observers independently coded 15 video-recorded Strange Situations. Mean reliability for the interactive scales was .78. Intercoder agreement for classification in subcategories was 93%; for classification in the three main categories, it was 100%.

Richters, Waters and Vaughn (1988) suggested that mother–infant relationships be empirically classified based on interactive behavior and crying during reunion episodes, through the use of two discriminant functions. The first function distinguishes B from non-B classifications, so that the discriminant function scores can be used in analyses that call for a continuous security variable. If the first function classifies the mother–infant relationship as non-B, then the second function is applied to decide whether the classification is A or C. When the 39 mother–infant dyads were classified in this way, 85% of the sample received the same classification determined by the observers.

Three attachment behavior variables were also computed. Scores on proximity seeking and contact maintenance in the two reunion episodes ($M$ correlation = .66) were summated, as were resistance scores ($M$ correlation = .58) and avoidance scores ($M$ correlation = .66).

Responsiveness. This has been defined as the mother’s ability to perceive and interpret accurately signals implicit in the infant’s behavior and to respond to them promptly and appropriately (Ainsworth et al., 1974). Mothers were asked to play with their infant in our laboratory playroom for 8 min before the Strange Situation procedure started. They were asked to behave as they were used to at home. Sessions were recorded on videotape; coders made notes of their observations, and the Ainsworth rating scale for measuring responsiveness (see Ainsworth et al., 1974 for details) was used to rate maternal behavior. Coders were, of course, unaware of subjects’ scores on the other variables. Intercoder reliability over 10 sessions was .92. The mean score on
the scale was 6.5 ($SD = 1.7$). A high score on this scale indicates a high degree of responsiveness.

Results

Concurrent Validity

The Strange Situation and the Attachment Q-set both were developed to measure attachment security. Their outcomes should, therefore, converge. The most simple and straightforward test of concurrent validity is to compare security scores (as measured through the Attachment Q-set) of the anxious and the secure subgroups (as classified according to Strange Situation criteria). The mean security scores and standard deviations of A/C and B children are presented in Table 1. Secure children tended to score somewhat lower on security as measured through the Attachment Q-set; they appeared, therefore, to be more secure. The analysis of variance (ANOVA), however, did not show a significant effect. Introducing desirability as covariate, and sex of child and attachment classification as factors did not produce a significant effect either, $F(1, 34) = 1.96$, $ns$. The same analyses were executed with computed classifications, according to the algorithm of Richters et al. (1988). The results were comparable. The ANOVA with security means (as measured through the Attachment Q-set) and computed attachment classifications did not show a significant effect, $F(1, 37) = 2.38$, $ns$, not even when desirability was introduced as a covariate and sex of child and attachment classifications were included as factors, $F(1, 34) = 1.90$, $ns$.

We did not find any significant (partial) correlations between security as measured through the Attachment Q-set and resistance, avoidance, and proximity seeking/contact maintaining as measured in the Strange Situation. Bivariate correlations with proximity seeking/contact maintaining, resistance, and avoidance were $-.08$, $.20$, and $.10$, respectively. Partialling out desirability, we found correlations of $.22$, $.16$, and $.10$, respectively. Security as

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means and Standard Deviations of Security and Responsiveness Scores in the Secure and Anxious Subgroups</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>$n$</th>
<th>Security</th>
<th></th>
<th></th>
<th>Responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxious</td>
<td>18</td>
<td>-.377</td>
<td>-.159</td>
<td>6.06</td>
<td>1.76</td>
</tr>
<tr>
<td>Secure</td>
<td>21</td>
<td>-.429</td>
<td>-.112</td>
<td>6.90</td>
<td>1.58</td>
</tr>
</tbody>
</table>

$^aF(1, 37) = 1.4$, $ns$. $^bF(37) = -1.59$, $p = .06$ (one-tailed test).
measured through the Attachment Q-set correlated .36 ($p = .012$) with security as determined by the first discriminant function.

**Predictive validity**

**Responsiveness.** Mothers in the secure subgroup tended to receive higher responsiveness scores than mothers in the anxious subgroup. Bivariate correlations between responsiveness and attachment behaviors—proximity seeking/contact maintaining, resistance, and avoidance—were not significant, but controlling for ICQ temperament variables resulted in a significant partial correlation between responsiveness and resistance ($r_p = - .37, p_1 = .015$). Mothers who showed more responsiveness in the free-play situation had children who showed less resistant behavior in the Strange Situation. No significant partial correlations were found for the other attachment behaviors.

The bivariate correlation between responsiveness and security as measured through the Attachment Q-set was not significant ($r = - .15; p_1 = .17$). Partial correlations controlling for desirability and for temperament variables, respectively, were $- .12$ (ns) and $- .20$ (ns). There were no significant correlations between responsiveness and outcome of the Attachment Q-set.

**Temperament.** No significant differences between the anxious and secure subgroups were found on the ICQ temperament variables: difficultness, adaptability, cuddly, persistence, and mood. ANOVAs with sex of child as a factor and socioeconomic status or responsiveness as covariates did not yield significant effects for attachment classification. Taking into account Belsky and Rovine’s (1987) hypothesis, we divided the children in two groups: A1 to B2, and B3 to C2. Analyses of covariance with the temperament variables did not produce any significant result. We found only one significant correlation for attachment behaviors: Persistence correlated with resistance ($r = .33, p = .04$). The more persistent and disobedient the child was, the more resistant behavior he or she showed during the Strange Situation. This correlation, however, was the only significant one out of 15 computed (5 temperament variables $\times$ 3 attachment behaviors).

In Table 2, bivariate correlations between temperament variables and security, dependency, and sociability as measured through the Attachment Q-Sort are presented. Partial correlations, controlling for desirability, were also computed. Because higher security scores indicate more anxious attachment feelings, an unexpected, significantly negative bivariate correlation was found between difficultness and security. The more difficult a child was, the more secure he or she was rated. An unexpected high negative partial correlation was also found between adaptability and security. The less adaptable to new circumstances a child was, the more secure he or she was rated with the
TABLE 2
Bivariate and Partial Correlations (Partialling Out Desirability) Between Temperament Variables and Security, Dependency, and Sociability as Measured Through the Attachment Q-set

<table>
<thead>
<tr>
<th>Temperament variables</th>
<th>Security</th>
<th>Dependency</th>
<th>Sociability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bivariate</td>
<td>Partial</td>
<td>Bivariate</td>
</tr>
<tr>
<td>Difficultness</td>
<td>-.31*</td>
<td>-.10</td>
<td>-.42*</td>
</tr>
<tr>
<td>Adaptability</td>
<td>-.02</td>
<td>-.69***</td>
<td>-.59***</td>
</tr>
<tr>
<td>Cuddly</td>
<td>-.23</td>
<td>-.33*</td>
<td>-.25</td>
</tr>
<tr>
<td>Mood</td>
<td>.03</td>
<td>-.25</td>
<td>-.35*</td>
</tr>
<tr>
<td>Persistence</td>
<td>.15</td>
<td>.11</td>
<td>.04</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.

Q-set procedure. Furthermore, the less a child liked to be cuddled, the more insecure he or she was rated.

Bivariate and partial correlations between dependency and temperament variables were in the expected direction. More difficult, less adaptable, and less cheerful children were rated lower on dependency; that is, they were seen as more dependent. Children who liked to be cuddled were also seen as more dependent. Only one (bivariate) correlation with sociability appeared to be significant: As expected, more adaptable children were rated as more sociable on the Attachment Q-set. In sum, dependency appeared to tap some temperamental aspects of children's behavior.

Against our expectations, however, security also correlated with one of the most important temperament dimensions: difficultness. More difficult children appeared to be rated as more secure. Furthermore, when controlling for desirability, we found that less adaptable children were rated as more secure.

Some caution is needed, however, in interpreting partial correlations deviating strongly from their bivariate equivalents. The status of the desirability measure is not quite clear because independent validation of this aspect of the Attachment Q-set has not been carried out yet.

Discussion

Although secure children scored somewhat lower on Attachment Q-set (in-) security than anxious children, this trend in the expected direction was by no means significant. Attachment Q-set security was significantly correlated with security as computed according to the Richters et al. (1988) algorithm, however. The correlation implies about 13% shared variance, which seems mod-
est from the perspective of converging measures. We are therefore not able to establish an impressive concurrent validity figure for the Attachment Q-set.

Furthermore, security as measured through the Attachment Q-set was not related to caregivers' responsiveness, although attachment theory suggests there should be some correlation between the way in which caregivers react to their children's signals and the quality of their attachment relationship. This result converges with the outcomes of another study on the validity of the Attachment Q-set, carried out with a sample of 48 infants (Van der Pol, 1987). In that study, responsiveness was measured through an event-sampling method applied in a natural setting.

Finally, the correlations between the temperament dimensions of difficultness and adaptability and security as measured through the Attachment Q-set were contrary to our expectation. In attachment theory, it is not only doubted that temperament has any direct link with attachment security, but it also would not have been predicted that secure children would be more difficult or less adaptable to new circumstances than would anxious children (Bretherton, 1985). We are therefore not able to establish the predictive validity of this new attachment security measure.

It is not possible, however, to evaluate the validity of the new instrument in a definitive way. First, the main problem of validity studies is, of course, the validity of the criteria against which a new measure is evaluated. In our study, responsiveness appeared to be related to anxious and secure attachment, as measured through the Strange Situation, but the relation was only marginally significant. We discovered, however, a rather strong and significant partial correlation between responsiveness and resistance (one of the most important indicators of anxious attachment). Therefore, we may have some confidence in the validity of these measures.

Second, we restricted ourselves to the parental version of the Attachment Q-set. It seems somewhat paradoxical to expect insensitive parents to observe their children's subtle avoidance and resistance in natural settings when it is already rather difficult for trained observers to register those attachment behaviors in the stressful Strange Situation. The observer's version of the Attachment Q-set may be more apt to tap the security dimension. Nevertheless, controlling for differences in maternal responsiveness did not lead to a significant effect of attachment classification on Q-set security.

Third, we did use the latest published version of the Attachment Q-set, but we did not take into account recent revisions (Waters, 1987), involving reformulations of about half the items. For this revised version, which has not been published yet, the criterion sorts (Waters & Deane, 1985) are no longer adequate representations of security, dependency, and sociability.

Fourth, our sample size was not very large, and the power of the statistical tests was therefore restricted. The possibility cannot be excluded that in
a larger sample our null hypothesis of no relation between Strange Situation outcome and attachment Q-set would be rejected.

Fifth, we carried out the study in a European country, thus possibly contaminating internal validity issues with external or cross-cultural validity issues. Elsewhere, however, we have presented some evidence for the cross-cultural validity of the Strange Situation (Van IJzendoorn & Kroonenberg, 1988). An alternative measure for attachment security should also be applicable across (at least Western) cultures.

Further research on the validity of the Attachment Q-set is necessary to determine the validity of this new instrument, which would greatly enhance researchers' opportunities to study attachment in larger groups and across broader age ranges. We believe that the Q-set and the Strange Situation are important tools for examining the development of parent–child relationships because they tap somewhat different dimensions or aspects of attachment. Considerable covariance between the two measures should therefore not be expected, and other ways of establishing the Q-set's validity must be explored, for example, determining different antecedents and consequences of different Attachment Q-set scores.

REFERENCES


ment as a potential developmental influence on attachment. In J. V. Lerner &
R. M. Lerner (Eds.), Temperament and social interaction in infants and children.
in the Netherlands. Paper presented at the fourth biennial International Conference
on Infant Studies, New York.
Kohnstamm, G. A. (1986) (Ed.). Temperament discussed: Temperament and devel-
opment in infancy and childhood. Lisse: Swets & Zeitlinger.
mother attachment: The origins and developmental significance of individual dif-
infant–mother relationships from interactive behavior and crying during reunion.
Tavecchio, L. W. C., & Van IJzendoorn, M. H. (1987). Perceived security and ex-
tension of the child’s rearing context: A parent-report approach. In L. W. C. Tav-
eccio & M. H. van IJzendoorn (Eds.), Attachment in social networks: Contributions
to the Bowlby–Ainsworth attachment theory (pp. 35–92). Amsterdam:
Thompson, R. A., & Lamb, M. E. (August, 1982). Temperamental influences on
stranger sociability and the security of attachment. Paper presented to the Ameri-
can Psychological Association, Washington, D.C.
Waters en Deane [Measuring attachment through Waters and Deane’s Q-sort]. Un-
Van IJzendoorn, M. H., & Kroonenberg, P. M. (1988). Cross-cultural patterns of
attachment: A meta-analysis of the Strange Situation. Child Development, 59,
147–156.
Van IJzendoorn, M. H., & Tavecchio, L. W. C. (1987). The development of attach-
ment theory as a Lakatosian research program: Philosophical and methodological
aspects. In L. W. C. Tavecchio & M. H. van IJzendoorn (Eds.), Attachment in
social networks: Contributions to the Bowlby–Ainsworth attachment theory (pp.
Waters, E. (1987, April). Open workshop on Q-sort methods for assessing attach-
ment behavior. Open workshop organized at the biennial meeting of the Society
for Research in Child Development, Baltimore, Maryland.
Waters, E., & Deane, K. E. (1985). Defining and assessing individual differences in
attachment relationships: Q-methodology and the organisation of behavior in in-
fancy and early childhood. In I. Bretherton & E. Waters (Eds.), Growing points
of attachment theory and research (pp. 41–65). Monographs of the Society for
Research in Child Development, 50, (1–2, Serial No. 209).

Received February 8, 1988