3

The Development and Validation of the AAP

This present chapter describes the development and validation of the Adult Attachment Projective Picture System. This discussion precedes a discussion of the details of the coding and classification system (see Chapters 4 and 5) to provide the reader with basic overarching information about the AAP.

The core of representational assessments is the view that narrative descriptions of experience are individuals' subjective constructions of their lives guided by internalized mental representations. In the course of development, and particularly in instances of compromised care, the personal and affectively charged elements in these "stories" are distorted in attempts to keep certain aspects of attachment-related distress carefully deflected. Experience and affect are transformed in order to be acceptable to consciousness and sometimes "locked away" and defensively excluded from conscious attention and memory. We discuss an attachment theory approach to defensive exclusion in Chapter 5. The point for the present discussion is that the goal of representation-based assessment is to uncover distortions and unlock walled-off elements in order to see through individuals' carefully constructed defensive maneuvers and thereby render the themes contained within their stories amenable to interpretation.

Historically, although attachment theory is grounded in naturalis-

tic observation (Ainsworth, 1964; Bowlby, 1951), some of the earliest assessments of attachment were representational. Bowlby was intrigued by the notion of using projective methodology to examine children's responses to separation and loss (Bowlby, 1973; Klagsbrun & Bowlby, 1976). Since attachment theory's "move to the level of representation" (Main et al., 1985) in the 1980s, researchers have developed interpretative schemes that draw upon the individual's ability to use symbolic representation and to organize knowledge conceptually (Bretherton, 2005; Waters & Waters, 2006). These mentalizing elements can be elicited from representation-based assessments beginning in the preschool years. Many representational assessments for children follow semiprojective methodologies for which researchers have developed rating scales for attachment-relevant constructs (e.g., sensitivity) or classification schemes (secure, avoidant, ambivalent, disorganized). The assessment stimuli include pictures or story stems that the children enact and narrate using family dolls. Detailed discussion of these forms of assessment for children is beyond the scope of this volume. The reader is referred to Solomon and George (2008) for a comprehensive review of representational and behavioral child attachment measures.

Representational assessment of adult attachment originated with the AAI (George et al., 1984/1985/1996), which is considered to be the "gold standard" measure for researchers and clinicians following the developmental attachment tradition. The AAI is a quasi-clinical interview during which individuals describe their childhood experiences with attachment figures using a series of standardized questions and probes that are designed to elicit memories of attachment figures in contexts that generate attachment distress, including separation, physical and psychological hurt, rejection, loss, and abuse. Individuals' accounts of experience vary in the extent to which their stories reveal unity or coherence among the network of attachment memories. Each interview is examined for one of three primary attachment patterns as analogous to the infant attachment groups—secure/autonomous, dismissing, and preoccupied—and for evidence of an unresolved state of mind regarding loss through death or abuse. The structured open-ended questioning of attachment-oriented themes has been demonstrated to be a valuable design feature of the AAI. This form of questioning encourages individuals to create their life story in a "conversation" with the interviewer that provides a picture of childhood attachment experiences and the meanings that surround these experiences. Readers interested in learning more about the AAI are referred to Hesse's (2008) discussion of AAI development, validation, and use.

Unlike the AAI, the AAP was not designed to elicit autobiographical narrative; rather, it systematically activates attachment by presenting a standardized set of projective attachment stimuli. Individuals are asked to create narrative "stories" about drawings that depict attachment scenes. Like the AAI, the responses to the AAP stimuli are conceived as the product of individuals' internal working models of attachment.

Our picture system approach follows the semiprojective assessment already in use the field to assess attachment in children, in particular the Attachment Doll Play Assessment (Solomon et al., 1995). The common denominator of projective methodology (i.e., free-response task) is the access to conscious and unconscious thoughts and emotions through verbal responses to a standardized set of ambiguous visual stimuli (Hilsenroth, 2004). The administration technique is unstructured. Unencumbered by administrative directives, individuals are encouraged to respond freely, guided only by a few basic standardized, open-ended questions. As such, projective methodology is not contaminated by the self-serving biases that plague self-report measures, have a lower risk of the exaggerations and minimizations of experience found in clinical interviews, and are economical and easy to use (Hilsenroth, 2004). The narratives provide a rich picture of interpersonal and behavioral dimensions, revealing patterns of unconscious and automatic defensive processing that lead to reliable interpretation when interpretive rules follow standardized guidelines (Leichtman, 2004). Although criticized by some assessment experts as subject to interpretive bias, and poor validity and reliability, large-scale studies of reliability (interrater, test–retest reliability) and predictive validity demonstrate that the projective technique is a valid form of assessment (Hilsenroth, 2004; Meyer, 2004; Wood, Nezworski, & Stejskal, 1996), especially when interpretation has a strong theoretical foundation (McClelland, Atkinson, Clark, & Lowell, 1953).

PICTURE STIMULI SELECTION PROCESS

The scenes that comprise the stimuli for the AAP were selected to capture three core features of attachment theory as defined by the Bowlby–Ainsworth model. The first feature is observing attachment under conditions that activate the attachment system. Bowlby's seminal attachment trilogy (1969/1982, 1973, 1980) stressed the importance of observing attachment in contexts that threaten or compromise physical or psychological safety. Strictly speaking, of course, the internal working model of attachment is not directly observable, and assessment must activate

the system in order to "see" the variations in its representational manifestations. A study of the "contents" and vicissitudes of attachment representation is inferred directly during assessment. The importance of activating attachment during assessment is somewhat controversial. Bowlby clearly described its importance; Ainsworth integrated these ideas directly into her naturalistic observation format and designed the Strange Situation to be a mini-drama of attachment activating events (Ainsworth, 1967; Ainsworth et al., 1978). Recently, the question of how to assess attachment in children based on scripts and in adults based on measuring conscious social cognitions has raised the issue of how to interpret attachment "data" when the system is not activated. The recent script literature acknowledges the overlaps among attachment group distinctions, for example, when individuals know the scripts of security but may not themselves be secure (Waters & Waters, 2006). More problematic is evaluating attachment groups or dimensions based on social cognitive responses to generic questionnaire items that bear only semantic resemblance to actual activating events (Maier, Bernier, Pekrun, Zimmerman, & Grossmann, 2004). Therefore, in developing the AAP stimulus set, we followed Bowlby's model and selected scenes that portray major attachment life events—separation, solitude, fear, and death.

The second feature is the availability of an attachment figure. This feature is one of the defining constructs in attachment theory (Ainsworth et al., 1978; Bowlby, 1969/1982). Ainsworth showed that attachment security was related to the attachment figure's direct availability and responsiveness. Compromises in availability are associated with the development of avoidant and ambivalent-resistant forms of organized insecure attachment. Unable to achieve proximity successfully through direct signaling, these insecure children must rely on alternative mechanisms to achieve proximity and the caregiver's attention when distant or distracted (George & Solomon, 2008; Main, 1990; Solomon & George, 1996). The hallmark of disorganized attachment is the child's subjective experience of attachment figure abandonment and unavailability (George & Solomon, 2008; Solomon et al., 1995). Solomon and George (2000) developed the concept of "abdication of the caregiving system" to describe the caregiving failures that leave the child feeling helpless and vulnerable and having to take responsibility for his own attachment needs.

West and Sheldon-Keller (1994), expanding on Weiss (1982), pointed out that there are qualitative differences over the lifespan in the form of the kind of attachment figure availability required to assuage distress.

In infancy and early childhood, attachment figures need to be physically present and accessible. By adolescence and throughout adulthood, individuals predominantly use "psychological" or representational proximity to replace in many situations the physical access to attachment figures required by children (see also Allen, 2008; Main et al., 1985).

Following this theoretical foundation, the AAP picture stimuli depict attachment figure availability in two forms. One form explicitly portrays the proximate availability of an attachment figure by drawing characters in dyadic pairs. We call these stimuli the dyadic pictures. The other form portrays individuals alone. We call these stimuli the alone pictures. In the absence of a visible cue, availability must be created by describing an internalized or physically present attachment figure, a task consonant with the abstract thought that characterizes adult mental representations of attachment (see Main et al., 1985). We define the attachment figures created in alone responses in terms of the *internalized secure base* and *haven of safety*, respectively (see Chapter 4).

The third feature captures attachment theory's lifespan view. Bowlby and Ainsworth laid the foundation for thinking about attachment contributions to development and mental ill-health beyond infancy (Ainsworth, 1989; Bowlby, 1969/1982). We integrated this lifespan view in the AAP stimulus set by including characters that depict childhood to old age.

The drawings that comprise the AAP picture system were originally selected based on pictures in children's story books and coffee-table books thought to depict attachment situations. A large set of drawings was developed and presented to undergraduate college students to rate for attachment pull (see West & Sheldon-Keller, 1994, for a detailed description of the stimulus selection process). The initial AAP stimulus set was composed of three scenes. These scenes depicted solitude (*Child at Window*), nighttime separation (*Bed*), and separation in the context of illness or death (*Ambulance*). Other scenes were later added to the set to enhance the AAP's ability to activate attachment by presenting a broader spectrum of attachment events.

THE TASK

The AAP stimulus set is comprised of eight line drawings. The first stimulus serves as a warm-up for the AAP task; it is not an attachment scene. The seven other stimuli depict theoretically derived attachment scenes. The drawings contain only sufficient detail to identify an event; facial

expressions and other potentially biasing details are absent. The characters were drawn to capture a diverse range in cultural background, gender, and age. The AAP stimuli are presented in Figures 3.1–3.8. The scenes include *Neutral*—two children play with a ball; *Child at Window (abbreviated as Window)*—a child looks out a window; *Departure*—an adult man and woman stand facing each other with suitcases positioned nearby; *Bench*—a youth sits alone on a bench; *Bed*—a child and woman sit opposite each other on the child's bed; *Ambulance*—a woman and a child watch ambulance workers load a stretcher into an ambulance; *Cemetery*—a man stands by a gravesite headstone; and *Child in Corner (abbreviated as Corner)*—a child stands askance in a corner.

The AAP is administered in a private setting. It can be used alone or as part of an assessment battery. If used with other assessments, the rule of thumb for research and clinical work is to administer developmental attachment measures before any other assessments that day. We often use the AAI in our work; therefore, we needed to know whether there was an administration effect for these two measures. We found no order effects in a subsample in which we counterbalanced the administration of the AAP and the AAI in the same session. Our own preference when these measures are administered in tandem is to give the long and ardu-
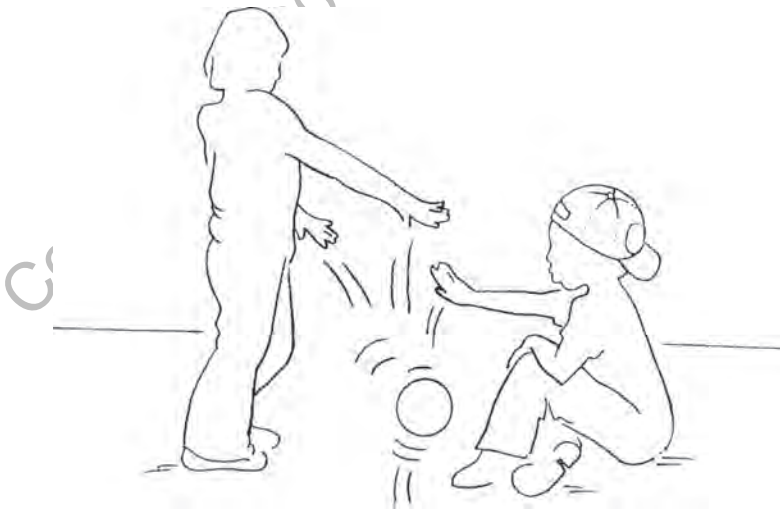
FIGURE 3.1. *Neutral*. Copyright 1997 by Carol George, Malcolm L. West, and Odette Pettem. Reprinted by permission.
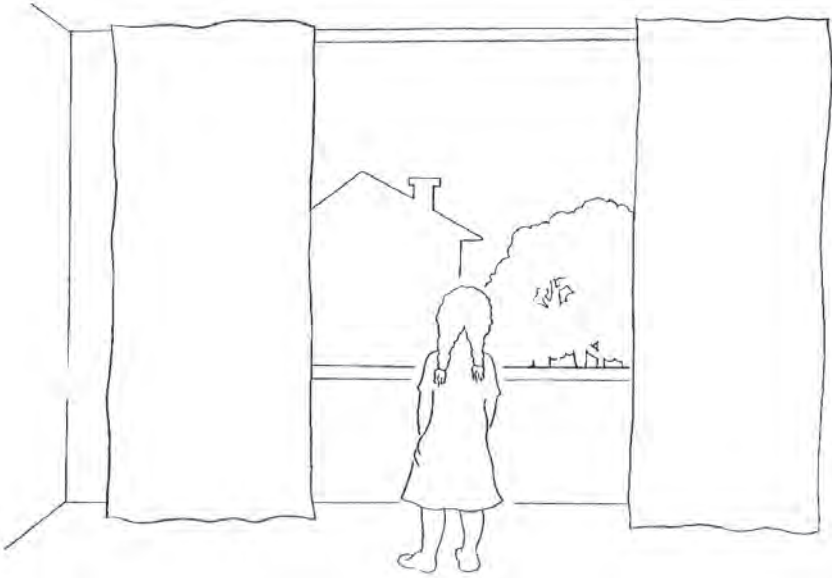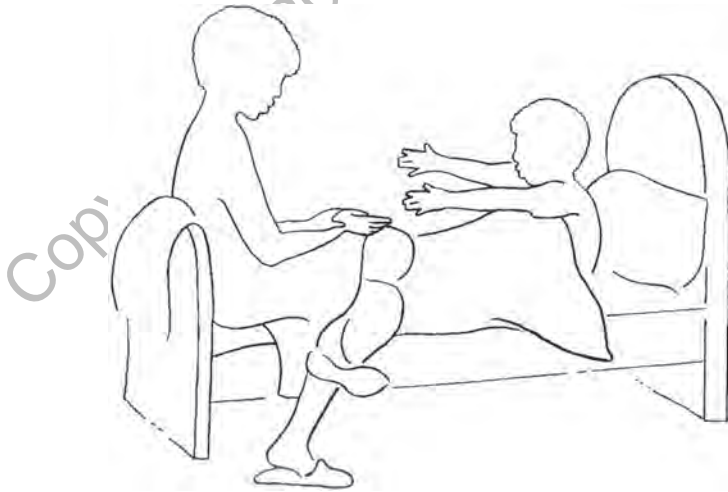
FIGURE 3.2. *Child at Window.* Copyright 1997 by Carol George, Malcolm L. West, and Odette Pettem. Reprinted by permission.

FIGURE 3.3. *Departure.* Copyright 1997 by Carol George, Malcolm L. West, and Odette Pettem. Reprinted by permission.

FIGURE 3.4. *Bench*. Copyright 1997 by Carol George, Malcolm L. West, and Odette Pettem. Reprinted by permission.

FIGURE 3.5. *Bed*. Copyright 1997 by Carol George, Malcolm L. West, and Odette Pettem. Reprinted by permission.

FIGURE 3.6. *Ambulance.* Copyright 1997 by Carol George, Malcolm L. West, and Odette Pettem. Reprinted by permission.
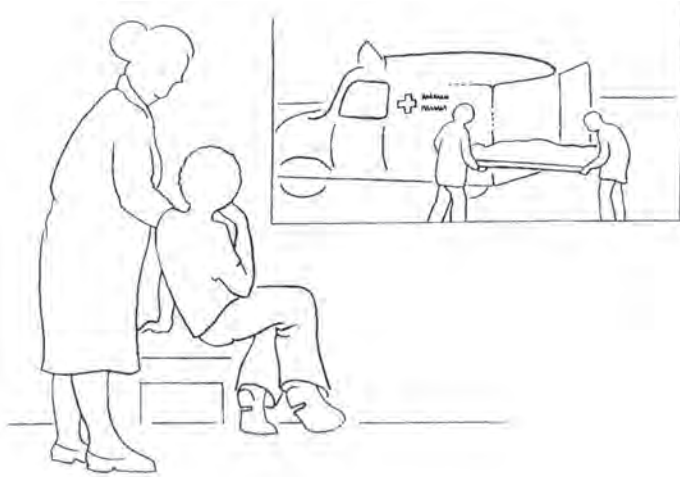

FIGURE 3.7. *Cemetery.* Copyright 1997 by Carol George, Malcolm L. West, and Odette Pettem. Reprinted by permission.

FIGURE 3.8. *Child in Corner.* Copyright 1997 by Carol George, Malcolm L. West, and Odette Pettem. Reprinted by permission.

ous AAI first. We find that the shorter and simpler AAP is a good way to "wrap up" the attachment portion of an assessment battery.

The AAP administration method combines apperceptive projective free-response and semistructured interview techniques. This approach has strong demonstrated success in adult and child attachment assessment (e.g., Bretherton, Ridgeway, & Cassidy, 1990; George & Solomon, 2008; Solomon et al., 1995). The interviewee is seated across from the administrator. The interviewee is handed each stimulus to hold while responding. The task begins with the *Neutral* warm-up scene.[1] The administrator begins by stating the instructions: "Describe what is happening in the picture, what led up to the events, what the characters are thinking or feeling, and what will happen next." One need not look far for the implications of the task instructions. At once, the individual is confronted with a paradox. The stimuli are clearly pictorial and at one level define the task as a purely perceptual one. The instructions simultaneously encourage a subjective response in that they require individuals to use their imagination as well as conscious and unconscious memories and mental concepts of attachment. Performance anxieties are eased by

---

[1] Variations of administration using more than one warm-up stimulus are described in Chapter 10.

assuring the interviewee that there are no incorrect answers. Throughout the task, the administrator waits comfortably for the interviewee to complete the response, using the above questions to prompt storytelling as needed. Once a response is completed, the administrator hands the interviewee the next picture stimulus and proceeds similarly through the remainder of the picture set.

The AAP is administered only using the full set of picture stimuli in the order designated above. Unlike some free-response assessments, the administrator may not select stimuli or alter the presentation order. We designed the presentation order so that less distressing stimuli are presented early on during the AAP, progressing to increasingly threatening attachment scenes. This design order parallels the design of other developmental attachment assessments (Strange Situation, doll play story stems, AAI).

Our conceptualization of the the AAP stress activation progression is supported in a neuroimaging study that examined attachment activation in the fMRI scanner. Buchheim, Erk et al. (2006; see also Chapter 10) reported increased activation of the right inferior frontal cortex over the course of the AAP task, an area of the brain that is associated with the process of suppressing unwanted emotion and reappraising unwanted emotions in unemotional terms. This study also found increasing activation in medial temporal regions of the brain (amygdala, hippocampus) for individuals with unresolved attachment. In summary, these results point to increased involvement of emotion control processes as the attachment stimuli portray increasingly threatening events.[2]

The AAP is tape-recorded and coded later from a verbatim transcript. The standard administration time is approximately 25 minutes. AAP transcripts are typically two to three pages long, sometimes longer in the case of clinical clients. Coding and classification is done by a trained reliable judge and typically takes between one-half and 2 hours.

Administrators do not need to have a background in attachment theory, assessment technique, or the AAP coding and classification system. The AAP has been administered by a range of different individuals, including women, men, undergraduate research assistants, experienced researchers, and clinicians, and has been administered in

---

[2] An interesting observation related to presentation order is when individuals' responses to the initial AAP stimuli include dysregulated attachment themes. This is an atypical response that we interpret as individuals' hypersensitivity to events as highly threatening. This type of responsiveness appears more frequently in research with clinical samples and in psychotherapy clients than in typical samples.

languages other than English (e.g., French, German, Japanese, Italian). Administration training typically requires three to four supervised practice cases. Ethical use of the AAP is important for the validity of the instrument and protection of research participants, clients, and patients. It is assumed, and we have indeed found, that most individuals who are asked to respond to the AAP have a positive attitude and are cooperative. Interviewees typically do not get upset during the AAP experience. No debriefing or follow-up is typically required. It is important to remember, however, that the pictures are powerful stimuli. In some situations, the administrator may detect some reluctance or distress in the individual's response to the task. Interviewers are trained to follow the probes carefully and not debrief or engage in caretaking during the task. Debriefing may be necessary if the interviewee is highly emotionally distressed during the task. On rare occasions, the interviewee asks to stop. We have developed clear guidelines to help administrators identify defensive resistances as compared to cues that would require terminating the administration session. If an interviewee is highly distressed, the AAP should be curtailed and re-administered on another day. This strategy is rarely needed with the AAP, which makes it amenable for use in stressed populations in which other attachment measures, especially the AAI, may be too stressful or prohibited (Buchheim & George, 2011; Szanjberg & George, 2011). These guidelines have been approved by national and international internal review boards that oversee research with human participants.

Training in the coding and classification of the AAP is also part of its ethical use. Evaluation to determine the individual's attachment response and representation cannot be done intuitively, even if one is trained in other attachment assessments. More information about training and use is available on the AAP website: *www.attachment projective.com*.

VALIDATION OF THE AAP

We have dedicated the past decade to validating the AAP. Beginning with Ainsworth's seminal work with the Strange Situation (Ainsworth et al., 1978), developmental attachment theory primarily uses a taxomic measurement approach; children and adults are placed into secure or insecure classification groups. This is not the only approach to measuring attachment, and some argue the merits of adding dimensional or other approaches to augment our understanding of the attachment con-

struct (Fraley & Spieker, 2003; Solomon & George, 2008; Solomon et al., 1995; Waters & Beauchaine, 2003). However, with the exception of the Attachment Q-sort methodology (van IJzendoorn, Vereijken, Bakermans-Kranenburg, & Riksen-Walraven, 2004; Waters, 1995; Waters & Deane, 1985), nontaxonomic approaches are poorly validated or define new constructs that are only tangentially related to the Bowlby–Ainsworth model of attachment (Crowell et al., 2008; Riggs et al., 2007). Therefore, we chose to follow the taxonomic tradition for validating the AAP and developed a classification scheme that discriminates among the four major adult attachment groups: secure, dismissing, preoccupied, and unresolved. Our approach led to the development of a set of unique assessment dimensions that permit examination of mechanisms that contribute to individual differences, which Waters and Beauchaine (2003) pointed out has been virtually ignored in the field of attachment. Recent research focusing on analyses of the AAP dimensions both supported the taxonomic classification approach and yielded new information about the underlying mechanisms of attachment. The "nuts and bolts" dimensions of our classification scheme are discussed in subsequent chapters, including the application of the AAP dimensions to case interpretation. Here we present our development and validation research that establishes the AAP as a valid taxonomic adult attachment classification measure.

We approached development and validation in three stages. The initial version of the AAP classification scheme was based on evaluations of 13 transcripts of men and women recruited from the community through newspaper advertisement. This work used a form of the AAP that included only six attachment stimuli; we had not yet developed the *Corner* scene. We examined the verbatim AAP transcripts of the attachment stimuli stories from a number of different aspects, including looking for attachment themes, specific content features, and descriptive images. This process led to developing a set of AAP content coding dimensions that we thought would differentiate among attachment group classifications. We also examined narrative discourse patterns, following the well-established AAI tradition. Our primary goal when we began this endeavor, and one that we maintained throughout our work, was to validate the AAP by establishing concurrent predictive validity for four attachment groups as designated by the AAI. Nine of the individuals in this initial group had completed the AAI prior to administration of the AAP. AAIs were classified blind by the first author. We checked our AAP classifications against the AAI classifications and then used our knowledge of the AAI classification to refine the AAP classification system

on a case-by-case basis. At this point, we added the *Corner* picture, a stimulus that extended the AAP attachment themes to include potential abuse[3] as well as coding rules for evaluating defensive processes.

The next step was to examine the validity of our classification scheme using two new samples and the eight-picture stimulus set. We continued to test for concurrent validity with the AAI and also sought to establish interjudge reliability. The first sample included a subsample of 25 mothers drawn randomly from a large ongoing study of infant risk conducted by Dr. Diane Benoit, which included a control sample of mothers of infants who were not at developmental risk. Dr. Benoit, a trained, reliable AAI judge without any knowledge of our AAP coding scheme, classified the AAIs and had members of her research team administer the AAP. The AAP development team (Drs. George, West, and Pettem) classified the AAP transcripts blind to AAI classification and infant risk status. The second sample included 24 women who had participated in a large-scale study of depression (West & George, 2002). AAIs were coded blind by the first author. Interjudge reliability was examined by comparing the independent classifications of the two other members of the development team, one of whom (Pettem) was blind to clinical group status.

The validity analyses based on these two samples demonstrated acceptable validity for the AAP. Interjudge reliability in this phase of developing the AAP ($n = 49$) was 87% agreement for four-group classifications (i.e., secure, dismissing, preoccupied, unresolved; kappa = .82, $p < .000$) and 97% agreement for secure versus insecure classifications (kappa = .73, $p < .000$). There was 92% AAP/AAI convergent agreement for four-group classifications (kappa = .89, $p < .000$) and 97% convergent agreement for secure versus insecure groups (kappa = .80, $p < .000$).

THE AAP VALIDATION STUDY

The third step in our validity research was to design a large scale psychometric investigation that added AAP test–retest reliability, discriminant validity, and AAI reliability to our research design.[4] This study consisted

---

[3] We are indebted to Dr. Diane Benoit for this suggestion.

of a sample of 144 participants represented by two subsamples of individuals, one from Calgary, Alberta, Canada (*n* = 73) and the other from northern California (*n* = 71). Adults between the ages of 18 and 65 were recruited from September 2002 to August 2003 from community and college settings using newspaper and Internet advertisement. Individuals with insufficient fluency in English were excluded from participation. Further exclusion criteria included a diagnosis of psychosis, or organic brain or central nervous system disorder. Three participants were judged Cannot Classify (CC) on the AAI and dropped from the validity sample. These exclusion criteria were necessary for this psychometric study because the attachment classifications derived from the AAI and the AAP are based largely on language discourse rating criteria that would be unduly biased by cognitive or psychiatric impairment or a poor ability to use the English language.[5] Dropping CCs from the analysis was necessary because it is atypical in community samples and there is no well-established understanding of this classification group (Hesse, 2008; we explore this issue further in Chapter 9).

One of the strengths of our validity sample was that it included both female and male participants, which adds significantly to the scarcity of adult attachment validation data for males. Of the 144 participants, 100 were female and 44 were male. The mean age for females was 36.2 years (*SD* = 15.2) and for males was 26.4 years (*SD* = 8.9). The males participating in this study were on average significantly younger than the females (*t* = 4.88; *p* < .000). There was a significantly greater number of Canadian female participants than U.S. women (*n* = 58) in the sample and there was a significantly greater number of U.S. men (*n* = 29) than Canadian men (*n* = 15) in the sample. The mean level of education was 14.7 years. The Alberta sample was uniformly Caucasian. The racial composition of the California sample was as follows: 49% Caucasian, 10% African American, 4% Hispanic, 12% Asian/Filipino, and 25% failed to specify.

Individuals participated in one-on-one sessions in a private office. Participants signed consent forms prior to beginning a session. During the session, research assistants first administered the attachment mea-

---

[5] The use of the AAP with individuals with cognitive and psychiatric impairment is being explored by researchers and clinicians. The first author is currently involved in research using the AAP with individuals with cognitive impairment. No results from this study are available at this time. Studies using the AAP with psychiatric patients have shown the AAP to be a valid measure for both out- and inpatients (Buchheim, Erk, et al., 2006; Buchheim & George, 2011)

sures, counterbalancing AAP/AAI administration in order to minimize potential administration order effects. A verbal intelligence measure was administered following the attachment measures. Other questionnaires were either filled out by participants at the end of the laboratory session or were mailed to the participants for completion in advance of the laboratory session. Approximately 12 weeks later (range 8–15 weeks), 69 individuals (48%) returned to complete the AAP a second time. Only the AAP was administered during the return testing session.

Measures

Adult Attachment Projective Picture System

The AAP picture system used in this study was the now standard eight-picture stimulus set. As in previous studies, AAPs were audiotaped and transcribed for later verbatim analysis. Three trained and reliable judges, blind to all information about the participants, independently coded the AAP transcripts. All transcripts were classified by at least two judges. Judge$_{AAP}$1 classified the entire set of AAPs; Judges$_{AAP}$ 2 and 3 classified subsets of AAPs that overlapped with Judge$_{AAP}$1. Judge$_{AAP}$ 2 classified 74 cases; Judge$_{AAP}$ 3 classified 135 cases.

Adult Attachment Interview

The most recent (third edition) version of the AAI (George, Kaplan, & Main, 1996) was used in this study. This version includes more questions and probes about trauma than in the earlier versions of the AAI. AAIs were classified blind by two reliable judges. The total number of AAIs was 130; 14 were dropped because of technical problems. Judge$_{AAI}$1 classified all of the AAIs in the sample; Judge$_{AAI}$2 classified 30 transcripts (n = 30; 21%) in order to establish interjudge reliability. AAI interjudge reliability was 85% agreement for four-group classifications (kappa = .72, p < .000; phi = 1.26, p < .000) and 87% agreement for secure versus insecure classifications (kappa = .63, p < .000; phi = .64, p < .000).

Balanced Inventory of Desirable Responding

The Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1998) is a 40-item social desirability assessment inventory that uses a 7-point Likert scale to assess two subscales: self-deception and impression management. The self-deception scale is designed to assess defensiveness toward personal threats and positively biased responding that the

respondent believes to be true (e.g., "I am a completely rational person"). The impression management scale is designed to measure responding that seeks to create a favorable impression on others (e.g., "I never take things that don't belong to me"). Each scale has 10 true-keyed and 10 false-keyed items. This measure has established acceptable validity (e.g., Holden, Starzyk, McLeod, & Edwards, 2000; Lanyon & Carle, 2007).

Verbal Intelligence

The vocabulary and similarities subtests from the Wechsler Adult Intelligence Scale—Revised (WAIS-R; Wechsler, 1981) were used to measure verbal intelligence. The verbal intelligence scales were selected from the WAIS because AAP responses and the coding system are language based.

Symptom Check List–90—Revised

The Symptom Check List–90—Revised (SCL-90-R; Derogatis & Cleary, 1977) is a validated and widely used measure of current levels of distress. The 90 questions yield separate scales related to nine symptom dimensions: somatization, obsessive–compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism. Participants are asked to rate the extent to which they felt distressed by the problem within the last 7 days on a 5-point scale ranging from extremely (0) to not at all (4). In addition to the symptom dimensions, the SCL-90-R also yields three composite scales: the Global Severity Index (GSI), an overall index of distress averaged across the nine dimensions; the Positive Symptom Total (PST), the number of symptoms endorsed; and the Positive Symptom Distress Index (PSDI), an index of severity of endorsed symptoms.

Results

Analyses of AAP and AAI classifications showed no administration order effects and we did no further analyses based on administration order. The AAP classification distribution of the 144 participants[6] was as follows: 25 (17%) were autonomous (F); 37 (26%) were classified as dismissing (Ds); 30 (21%) were preoccupied (E); and 52 (36%) were unresolved (U). This distribution was satisfactory for our study, which

---

[6] Not all measures were available for all participants.

sought to establish test characteristics and not population characteristics. The mean ages for each classification group were F, 28.27 years; Ds, 30.92 years; E, 38.57 years, and U, 34.12 years. Preoccupied participants were significantly older than autonomous participants ($F_{3,144} = 2.80$, $p < .05$). Age, gender, national residency, and years of education were not related to attachment classification.

Interjudge reliability for the AAP was calculated for both four-group (F, Ds, E, U) and two-group (secure vs. insecure) classifications. The classification matrices for the four-group and two-group classifications of Judge$_{AAP}$1 as compared with Judges$_{AAP}$2 and 3 are shown in Tables 3.1–3.4. Analyses using the kappa statistic demonstrated significant AAP interjudge reliability. There was 90% agreement between AAP Judges 1 and 2 on four-group classifications and 85% agreement between AAP Judges 1 and 3. There was a 99% and 92% respective concordance rate for two-group classifications for the two pairs of AAP judges.

TABLE 3.1. AAP Interjudge Classification Reliability between Judge$_{AAP}$1 and Judge$_{AAP}$2: Four Adult Attachment Classification Groups

Judge$_{AAP}$1	Judge$_{AAP}$2				
	F	Ds	E	U	Total
F	1	1	0	0	2
Ds	0	13	0	2	15
E	0	1	19	1	21
U	0	1	1	34	36
Total	1	16	20	37	74

Note. Kappa = .85, *p* < .000; phi = 1.41, *p* < .000.

TABLE 3.2. AAP Interjudge Classification Reliability between Judge$_{AAP}$1 and Judge$_{AAP}$3: Four Adult Attachment Classification Groups

Judge$_{AAP}$1	Judge$_{AAP}$3				
	F	Ds	E	U	Total
F	34	4	0	1	39
Ds	3	43	1	1	48
E	1	1	14	4	20
U	0	4	0	24	28
Total	38	52	15	30	135

Note. Kappa = .79, *p* <.000; phi = 1.38, *p* < .000.

TABLE 3.3. AAP Interjudge Classification Reliability between Judge$_{AAP}$1 and Judge$_{AAP}$2: Secure versus Insecure Adult Attachment Classification Groups

	Judge$_{AAP}$2		
Judge$_{AAP}$1	Secure	Insecure	Total
Secure	**1**	1	2
Insecure	0	**72**	72
Total	1	73	74

Note. Kappa = .66, *p* <.000; phi = .702, *p* < .000.

TABLE 3.4. AAP Interjudge Classification Reliability between Judge$_{AAP}$1 and Judge$_{AAP}$3: Secure versus Insecure Adult Attachment Classification Groups

	Judge$_{AAP}$3		
Judge$_{AAP}$1	Secure	Insecure	Total
Secure	**34**	5	39
Insecure	5	**91**	96
Total	39	96	135

Note. Kappa = .82, *p* < .000; phi = .82, *p* < .000.

Convergent predictive validity was calculated for the four-group and two-group classifications by comparing AAP and AAI classifications (*n* = 130). Classification disagreements among the three AAP judges and between the two AAI judges were independently resolved through consensus. Convergent agreement was examined using the kappa statistic. Analyses demonstrated significant convergent agreement between AAP and AAI (see Tables 3.5 and 3.6). Convergent agreement on AAP/AAI four-group classifications was 90%. Convergent agreement on AAP/AAI two-group classifications was 97%.

Test–retest reliability was determined for the 69 participants (48%; 39 females, 30 males) who returned for Time 2 testing. The kappa statistic was used to evaluate Time 1 and Time 2 AAP concordance for four- and two-group classifications. Test–retest reliability for four-group classifications was 84% and 91% for two-group classifications (see Tables 3.7 and 3.8). Stability from Time 1 to Time 2 was 82%, 96% for individuals judged dismissing, 62% for individuals judged preoccupied, and 80% for individuals judged unresolved. Of the 11 participants whose classifications changed over the retest period, seven were shifts between

TABLE 3.5. AAP/AAI Convergent Validity: Four Adult Attachment Classification Groups

AAP	AAI				
	F	Ds	E	U	Total
F	**19**	0	0	1	*20*
Ds	0	**30**	3	1	*34*
E	2	1	**26**	1	*30*
U	1	3	2	**40**	*46*
Total	*22*	*34*	*31*	*43*	*130*

Note. Kappa = .84, *p* <.000; phi = 1.49, *p* < .000.

TABLE 3.6. AAP/AAI Convergent Validity: Secure versus Insecure Adult Attachment Classification

AAP	AAI		
	Secure	Insecure	Total
Secure	**19**	1	*20*
Insecure	3	**107**	*110*
Total	*22*	*108*	*130*

Note. Kappa = .89, *p* < .000; phi = .89, *p* < .000.

TABLE 3.7. AAP Test–Retest Reliability: Four Adult Attachment Classification Groups

AAP	AAP retest				
	F	Ds	E	U	Total
F	**14**	1	0	2	*17*
Ds	0	**23**	1	0	*24*
E	0	2	**5**	1	*8*
U	3	0	1	**16**	*20*
Total	*17*	*26*	*7*	*19*	*69*

Note. Kappa = .78, *p* <.000; phi = 1.32, *p* < .000.

unresolved and organized classifications (four changed from unresolved to organized/resolved and three changed from organized/resolved to unresolved).

The discriminant validity analyses (MANOVAs) showed that AAP classifications were not influenced by verbal intelligence and social desirability. The results are presented in Table 3.9.

TABLE 3.8. AAP Test–Retest Reliability: Secure versus Insecure Adult Attachment Classification Groups

	AAP retest		
AAP	Secure	Insecure	Total
Secure	**14**	3	*17*
Insecure	3	**49**	*52*
Total	*17*	*52*	*69*

Note. Kappa = .77, *p* <.000; phi = .77, *p* < .000.

TABLE 3.9. Discriminant Validity: Verbal Intelligence and Social Desirability

	AAP classification group										
	F (*n* = 25)		Ds (*n* = 37)		E (*n* = 30)		U (*n* = 52)		*F*	*df*	*p*
	Mean	SD	Mean	SD	Mean	SD	Mean	SD			
WAIS Verbal	12.4	1.95	13.45	2.06	13.25	2.34	13.25	2.31	.86	3	.46
WAIS Similarities	11.50	1.97	13.00	2.63	13.43	2.52	13.00	2.68	2.05	3	.11
BIDR IM	7.45	3.58	6.95	3.89	8.73	4.10	8.73	4.14	1.89	3	.13
BIDR SDE	2.45	2.38	3.41	3.31	3.67	4.16	2.06	2.21	2.40	3	.07
BIDR Total	9.19	4.02	10.35	5.42	12.40	6.26	10.78	4.99	1.21	3	.31

We were also interested in examining differences among attachment groups on symptomology reported on the SCL-90-R GSI, PSDI, and PST. This measure is normed separately for women and men. The results were analyzed separately for gender using one-way analysis of variance. Analysis of PSDI scores for women was controlled for age because of a significant correlation between these two variables. We ran two sets of analyses. One compared the SCL-90-R scores of secure versus insecure adults; the other compared the scores of unresolved versus resolved/organized adults. Drawing from the attachment literature on risk (Cassidy & Shaver, 2008), we predicted that the scores for secure individuals would be lower than the scores for insecure individuals, and

the scores for unresolved individuals to be higher than resolved/organized individuals. The results of these analyses are shown in Table 3.10. The results showed no differences on SCL-90-R symptoms between women or men on either of these dimensions. These results suggest several things. One is that there were a lot of distressed women and men in this "community" sample. It is likely that many studies assume that these types of sample recruited from the general population are symptom free but, as Bakermans-Kranenburg and van IJzendoorn (2009) also emphasized, this assumption is not necessarily correct. We must stop and ask, "Who are the individuals who volunteer for a psychological study on 'relationships'?" We have found that it is not uncommon for individuals who volunteer for psychology studies to feel they have something to contribute or have "issues" they want to discuss. This may have led to an overabundance of insecure and troubled individuals in our sample, which was advertised as a "relationship" study. We present in Chapter

TABLE 3.10. SCL–90–R GSI, PDSI, and PST Scores for Women and Men

	GSI Mean (SD)	PDSI Mean (SD)	PST Mean (SD)
	Women		
Secure (n = 15)	54.48 (12.03)	55.55 (10.92)	54.95 (11.74)
Insecure (n = 82)	55.55 (10.92)	50.99 (7.92)	56.07 (10.32)
Organized (n = 63)	53.75 (9.64)	51.11 (7.82)	50.35 (7.37)
Unresolved (n = 34)	53.09 (8.52)	53.71 (10.22)	53.24 (8.74)
	Men		
Secure (n = 6)	59.67 (17.26)	50.83 (9.93)	62.17 (14.99)
Insecure (n = 36)	59.72 (13.19)	55.56 (9.36)	61.69 (9.74)
Organized (n = 25)	58.84 (14.78)	55.12 (9.94)	61.12 (10.74)
Unresolved (n = 17)	61.00 (11.96)	54.53 (9.02)	61.65 (10.24)

9 a new model of unresolved attachment that broadens the conceptualization of risk in attachment theory, following Bowlby's (1980) model of mourning. Utilizing this new approach in thinking about psychiatric risk samples may provide a better understanding of risk patterns in community and clinical samples than the standard comparisons we tested here.

OTHER AAP VALIDITY STUDIES

Concurrent Validity with the AAI in the German Language

Interjudge reliability and predictive concurrent AAP/AAI validity has been independently established for the AAP for use in the German language. This work also established the validity for using the AAP with psychiatric patients and to examine neurological patterns of attachment in a functional magnetic resonance imagery (fMRI) environment (Buchheim, Erk, et al., 2006; Chapter 10). The data comprise interviews from 74 females participating in several studies that included psychiatric patients and nonpatient controls (see Buchheim & George, 2011). Twenty-eight AAPs were administered in the fMRI environment (11 borderline inpatients, 17 controls), and 46 were administered using the procedure described above (5 controls, 21 borderline outpatients, 20 anxiety inpatients) and the AAI was administered 6 weeks later. Two reliable judges classified the AAPs, one using the German language transcripts and the other using English translations. The translator and the English-language judge were blind to all information about the participants. AAIs were classified by two reliable AAI judges who had established interjudge AAI reliability in other German samples. One AAI judge was blind to all information about the participants.

AAP interjudge reliability was reported only for four-group classifications. There was 98% agreement between the judges (kappa = .97, $p < .000$). AAP/AAI convergence for four classification groups was 84% (kappa = .71, $p < .000$), 91% for secure versus insecure (kappa = .70, $p < .000$), and 88% for unresolved versus resolved (kappa = .75, $p < .000$). These results establish the validity of the AAP in German-speaking samples for clinical and healthy participants. This reported AAP agreement is especially notable because the two AAP judges were coding and classifying independently in German and English.

Predictive Validity Studies

Independent studies using the AAP supports the validity of this measure. The AAP has been used in basic and clinical research, including studies of mothers' attachment in rel rising interest in and growing acceptance ofation to their children's adjustment and risk (Béliveau & Moss, 2005), adolescent development and maltreatment (Aikins, Howes, & Hamilton, 2009; Webster & Hackett, 2007, 2011; Webster, Hackett, & Joubert, 2009; Webster & Knoteck, 2007), foster care (Webster & Joubert, 2011), adult immigration (Van Ecke, 2006; Van Ecke, Chope, & Emmelkamp, 2005), depression (West & George, 2002), posttraumatic stress symptomology (Benoit, Bouthillier, Moss, Rousseau, & Brunet, 2010), and emotional development in psychiatric patients (Subic-Wrana, Beetz, Langenbach, Paulussen, & Beutel, 2007). Recent innovative research has used the AAP to examine the neurological and biochemical substrates of attachment in community and psychiatric samples (Buchheim, Erk, et al., 2008a; Buchheim, Erk et al., 2008b; Buchheim, Erk, et al., 2006; Buchheim & George, 2011; Buchheim et al., 2009; Warren et al., 2010; see Chapter 10). The AAP has been used to inform clinical practice (Finn, 2011; Lis, Mazzeschi, Di Riso, & Salcuni, 2011; Smith & George, in press) and child custody evaluation (George, Isaacs, & Marvin, 2011; Isaacs, George, & Marvin, 2009)

SUMMARY

We have been especially mindful of the issue of construct validity from the inception of the AAP. Our goal was to establish predictive concordant validity with the AAI, the only other well-validated developmental assessment of adult attachment in the Bowlby–Ainsworth tradition. Throughout the validation process, the AAP has demonstrated impressive agreement with the AAI, the gold-standard assessment in developmental adult attachment research. We can specify several reasons for this strong convergence. We briefly describe these reasons here, providing a bridge to describing the coding, classification, and application of the AAP that follows in subsequent chapters.

First and foremost, we emphasize that we did not start developing the coding and classification scheme "from scratch." We both had a strong foundation in attachment theory, and both authors had independently already developed and validated a number of other attachment assessment measures with other collaborators (e.g., AAI, George et al., 1984/1985/1996; Main & Goldwyn, 1985/1988/1994; Main et al.

2003; Attachment Doll Play Procedure, Solomon et al., 1995; Caregiving Interview, George & Solomon, 1996, 2008; Reciprocal Attachment Questionnaire, West, Sheldon, & Reiffer, 1987; Adolescent Attachment and Unresolved Attachment Questionnaires, West et al., 2000; West et al., 1998). We drew heavily from theory and integrated several ideas from our other measurement approaches. We found that adding the evaluation of defensive processes that George and Solomon originally conceived and validated for the analysis of their Attachment Doll Play Procedure (Solomon et al., 1995) and their maternal caregiving system assessment (George & Solomon, 1996, 2008) to be essential to our success in differentiating among the insecure adult classification groups (see Chapters 4 and 5 for discussions of defensive exclusion and AAP content coding).

Another reason for the strong AAP concurrent validity is that we were guided during our individual development phase by evidence from AAI research that had established validity for assessing attachment status based on narrative patterns. This led us to develop the AAP as a narrative-based assessment, rather than adopting some other format such as a questionnaire or a thematic approach as in the Rorschach. Interestingly, however, in contrast to the "sin qua non" role of narrative coherence for analyzing the AAI, narrative coherence does not contribute meaningful information toward classification. There are clearly differences in these two different kinds of narrative assessments. The AAI is an autobiographical task that depends first and foremost on the ability and willingness of the interviewee to engage in a cooperative conversation with the interviewer. Grice's philosophical narrative maxims, the fundamental guidelines Main used to develop the evaluation of coherence for the AAI, are conversation rules. The AAP is not an autobiographical task, and the structure does not take a conversational form. The administration instructions release the interviewee from the constraints of interpersonal discourse. Main's AAI interpretations of Grice's maxims added little to interpreting the AAP. But what was useful about AAI coherence in relation to the AAP was pointing out how narrative structure and content that heretofore was predominantly used only in clinical interpretation could be systematically described and examined empirically.

Rather than discourse coherence, the AAP assesses how the narrative response reveals *attachment coherence*, which we define as the representational integration of attachment and caregiving, that is, coherence in the relationship. *Attachment-coherent* doll-play stories characteristically portray children's needs clearly, attachment figures as sensitively attuned and protective, and an integrated, capable self, rather than restriction, confusion or a disorganized self (Solomon et al., 1995).

Descriptions of attuned and integrated responses combined with the mother's desire to achieve flexibility and balance between her caregiving system and her child's attachment needs are the hallmark of caregiving security (George & Solomon, 2008; Solomon & George, 1996).

These integrative elements in the narrative responses of adults and children provide the conceptual foundation for the evaluation of *attachment coherence* in the AAP, defined as the flexible integration of the attachment and caregiving systems and the portrayal of an autonomous and integrated self. The degree to which individuals achieve *attachment coherence* in the AAP depends on the shifting balance between adaptive processes and defensive exclusion in their efforts to give meaning to and find meaning in each picture stimulus.

This brings us back to defensive processes. Defense limits *attachment coherence*. George and Solomon's research has demonstrated that insecure attachment and caregiving (organized-insecure and disorganization) involves heightened defensive exclusion of the "normal" operation of the attachment system (George & Solomon, 1996, 2008; Solomon et al., 1995). Because the purpose of defensive exclusion is to suppress direct expression of attachment thoughts and feelings, assessment attends to defensive substitutions or what is unleashed when defensive exclusion breaks down. All of the features of attachment coherence (and incoherence) are described in the chapters that follow, delineated as we describe more fully the details of the AAP coding scheme and the AAP "meaning" associated with each of the different attachment groups.

Another factor that we believe contributed to the success of the AAP/AAI concordance rate is our restriction to developing a scheme that identifies only the major attachment categories (F, Ds, E, and U). The AAI classification scheme identifies subgroups within each major classification group (e.g., five forms of the secure classification group), analogous to Ainsworth's original group descriptions for the Strange Situation (Ainsworth et al., 1978). The field of attachment, however, rarely specifies category subgroups in research or clinical application, and in Chapter 10 we discuss how new classification groups and subgroups have emerged from empirical need rather than theoretical relevance. For most research purposes, the classification groups are collapsed for statistical analysis (i.e., secure vs. insecure; organized vs. disorganized). The taxonomic approach is based on identifying major groupings, a task that has been successfully established for the AAP through our validity work. Classification subgroups presumably provide some evidence of different underlying patterns within a particular classification group. But these

underlying patterns have never been clearly explicated in attachment theory, other than to say that different subgroup patterns exist.[7] The attachment meaning of some of these subgroups is poorly understood, such as the irrational fear of death associated with the dismissing classification group (Ds_4 subgroup on the AAI).

Finally, following Ainsworth's lead stressing the importance of "patterns of attachment" (Ainsworth et al., 1978), and the subsequent success of this approach in developing other attachment assessments, the AAP coding dimensions elucidate several different kinds of representational processes that we combine as patterns in order to derive a classification. Classification assignment follows explicitly schematized rules that define how to think about these patterns within a hierarchically integrated series of decision points.

The work that we described in this chapter establishes the AAP picture system as a valid measure for assessing adult attachment. Validation work is never finished. We hope that others will pick up this task in their research groups and clinical practice.

---

[7] We have had some success in identifying subgroups in our later work, but this has never been our goal. Rather, we have found that our defensive exclusion approach to classification is helpful in examining the underlying mechanisms, especially defensive processes that differentiate individuals among classification groups and individuals placed in the same attachment group.