

Final Copy

WORKING PAPER FOR THE MacARTHUR RESEARCH NETWORK ON  
ADOLESCENT DEVELOPMENT AND JUVENILE JUSTICE

**METHOD, MEASURES, AND PROCEDURES**  
**FOR THE**  
**JUVENILE ADJUDICATIVE COMPETENCE STUDY**

August, 2002

### *Introduction*

This paper reports the methods employed by the MacArthur ADJJ Research Network in its study of youths' adjudicative competence, performed in 1998-2002. It is written in a technical, archival style that is not intended for publication in professional journals. The paper was prepared to serve two main purposes.

First, the paper provides a detailed account of the design, method and procedures for the study, both for archival purposes and for use by individuals who have a need to know more than will appear in journal articles on the study. Few professional journals provide sufficient print space to describe research methods at a level of detail that is needed by persons who wish to scrutinize the limitations in a study's design or conduct, or who need sufficient detail to inform their own potential research on the same topic. This paper provides that greater detail, and it will be made available as a document on the Network's website.

Second, the paper is intended to provide a uniform reference for the Network's researchers as they prepare to describe the study in numerous publications for various audiences. Those publications will require selective reporting of the contents of this paper (due to print space limitations), and they will require somewhat different styles of reporting depending on the legal, scientific or lay audiences for whom the articles are being written. This working paper offers a master document from which information can be extracted and reported in various forms, while ensuring essential uniformity across publications regarding the facts of the study's design and methods.

### ***Overview of the Study's Purpose and Sampling Design***

The nature of the problem that this study addressed, and the theory and logic for its design, will be described in other documents available from the Network. A brief overview will be helpful, however, in orienting the reader to the nature of the study's design and methods. This is followed by sections describing the specifics of the study's design and method.

#### ***Purpose***

The Network wished to learn about the capacities of youths as defendants in their trials. Knowledge of youths' abilities to understand their trials, assist their attorneys, and make decisions in the trial process was needed in order to address a number of key issues in law, policy, and practice that have arisen in recent years regarding the adjudication of youths charged with offenses to be tried in juvenile or criminal court. Certain types of information were sought in order to assist in addressing the issue as it arises in application of the formal *legal criteria for competence to stand trial*. Other types of information were sought to address broader questions of youths' *judgment in the process of making decisions as defendants*, especially regarding the waiver of important rights of defendants (e.g., pleading).

The Network recognized that future applications of the information would require participants of various ages, gender, and ethnicities whose abilities could be compared in order to address critical questions in law, policy and practice. Questions pertaining to youths' competence to stand trial could most profitably be addressed by obtaining information about the abilities of youths actually involved in the juvenile justice system, as well as adults in the criminal justice system to whom youths' abilities could be compared. If age differences were found, however, the potential developmental meaning of those differences might be obscured

unless the study also included youths and adults who were not involved in the justice system. These general considerations produced the four-group design (adolescents and adults, within the justice system or the general community) that is reflected in the method that the study employed.

### *Age Ranges*

The study was restricted to participants who were 11 through 24 years of age. Youths 11-17 were labeled the “adolescent” or “youth” group and individuals 18-24 the “adult” group. The lower age was selected because delinquency charges are rare for youths below 11. Age 17 was chosen as the upper end of the adolescent group because it is the highest age for which most juvenile justice systems in the U.S. have jurisdiction. Ages 18 through 24 were chosen as the “adult” comparison group because a majority of individuals charged with crimes as adults fall within this age range. Therefore, they represent a population of “adults” to whom the legal system is most likely to compare adolescents in questions that arise about defendants’ competence to stand trial. Moreover, using this group of “young adults” produced a more conservative estimate (compared to the inclusion, for example, of presumably more mature 30- and 40-year-olds) of differences between adolescents and adults in the types of cognitive abilities and developmental characteristics that we wished to study.

For descriptive and analytic purposes, this study grouped youths in the adolescent age range into three age classes: 11-13, 14-15, and 16-17. The boundaries for these age classifications were decided during the process of data analysis, taking into account the age “breaks” that would best frame the results obtained when we examined the relation of age as a continuous variable to the dependent variables of greatest importance for the study.

### ***Detained and Community Groups***

The study included individuals grouped as “detained” or “community” participants. Detained participants were currently involved in the juvenile or criminal justice systems. They were residing in juvenile detention facilities or in adult jails at the time of their participation in the study, most of them awaiting trial, although some were recently adjudicated and awaiting placement. Community participants were individuals who did not meet criteria for prior involvement with the juvenile or criminal justice systems (that is, had never been held overnight or longer in a lock-up or detention center, and not currently charged with any offenses). Youths in the youngest age category were over-sampled in the juvenile detention centers in order to provide a sufficient number within this category for analyses related to age.

### ***Gender, Ethnicity, and Socioeconomic Classification***

Since a detained sample based on the male:female ratio within detention centers and jails would produce too few female participants for analyses employing gender as a variable, we decided to oversample females for the detained group

We determined to include participants in the Detained sample without participant selection criteria for ethnicity or socioeconomic classification. In order to assure adequate representation of ethnic minorities, however, we decided to select study locales that varied in ethnic composition of their populations.

To facilitate the development of Community samples that would approximate the Detained samples in gender, ethnicity and socioeconomic classification, we solicited participants in communities that were (a) in the same geographic areas as the detention center and jails, and (b) most likely to yield individuals whose ethnic and socioeconomic characteristics were similar

to those that became apparent for the Detained samples as data collection in detention centers and jails progressed. As described later in “Procedures,” sampling in those communities became gradually more selective as the study proceeded, in order to create age/gender/ethnicity cell sizes that would be fairly similar for the Detained and Community groups.

### *Sites*

Several data collection sites were required for the study, in order to (a) avoid results that might be biased by characteristics of detained youths or adults specific to any one jurisdiction, and (b) obtain samples that included adequate proportions for analyses by ethnic and socioeconomic classifications. The four study sites, described later in this report, included both urban and rural areas, and as well as sites in the Western, Eastern, and Southern U.S.

It was anticipated that three of the sites would contribute similar numbers of participants to the study, while a fourth site was expected to contribute a smaller number of participants because it was implemented several months after the start of the other three sites. Each site was expected to contribute an approximately equal number of participants to the Detained and Community samples, and all sites were expected to contribute participants to every age/gender/ethnicity cell in the design. We anticipated, however, that the sites would be disproportionate in their contributions to the samples with regard to ethnicity, which would vary naturally according to the ethnic composition of Detained youths and adults for sites in different geographic or urban/rural sectors of the U.S.

## *Measures*

The purpose of the study required an assessment of youths' and adults' abilities associated with the legal concept of adjudicative competence, as well as their performance in situations in which defendants are expected to make important decisions affecting their welfare. As described later, this led to the selection of a measure of basic abilities associated with competence to stand trial (the *MacArthur Competence Assessment Tool for Criminal Adjudication*) as well as the development of a new instrument (the *MacArthur Judgment Evaluation*) to assess decision making in various contexts as a defendant in a delinquency/criminal adjudicative process.

Fundamental to the perspective of this study was the notion that youths might have deficits in various abilities associated with decisionmaking in legal contexts due to their relative psychosocial immaturity. The study required, therefore, that we identify relevant dimensions of psychosocial maturity that might be relevant for decision making, and that we develop ways to measure those dimensions. Review and extensive consideration of a large number of possible dimensions, together with a review of the very modest body of research on adolescent decision making, led to the selection of three dimensions that, in theory, continue to mature throughout adolescence and have a potential influence on individuals' judgment when making decisions. These were *Future Time Perspective*, *Risk Perception*, and *Resistance to Peer Influence*.

As described in detail later, these dimensions were operationalized in the construction of two types of measures of psychosocial maturity for use in the present study. One set of measures assessed individuals' status on these dimensions in a generalized manner—that is, with reference to no specific decision making context. A second set of measures assessed

individuals' status on these same dimensions, but in the context of making decisions in police encounters, attorney encounters, and deciding how to plead in the adjudicative process.

Given this overview, the remainder of this document provides details regarding the following matters of method and procedure:

- Participants
- Measures for Independent Variables
- Measures for Dependent Variables
- Measures of Psychosocial Maturity
- Data Collection Procedures

### *Participants*

A total of 1432 individuals were interviewed. Data for 37 of these individuals were excluded from analysis, providing a final sample of 1393 participants. These exclusions occurred for either of two reasons: (a) more than 25% of the participant's data were incomplete, or (b) the participant obtained an IQ score below 60. The latter exclusion criterion was employed in order to produce a sample that would be comparable in IQ to samples published in the earlier norming and validation study for one of the present study's dependent measures of abilities associated with competence to stand trial (the *MacArthur Competence Assessment Tool – Criminal Adjudication*: Poythress et al., 1999).

As described later in "Procedures," it was not possible to develop a description of the populations from which the Detained youths and adults were obtained. Data collection was not consecutive for admissions to the detention and jail facilities, because it occurred only on days when the data collectors were present at the facilities. The project did not collect demographic information on all admissions to the facilities during the time of the study. An unknown number of detained youths and adults were "screened out" by staff or youth advocates (see "Human Subjects" below) based on their vulnerable conditions at the time of potential research

participation or their expressed disinterest in participating. Thus virtually all individuals who were approached by data collectors agreed to participate, and only two Detained youths discontinued participation partway through the procedure.

### ***Samples: Demographic Description***

The final sample of 1393 participants is described in **Table 1** according to age, gender, ethnicity and socioeconomic status. We obtained a similar number of Detained youths (n = 453) and Community youths (n = 474), and an equal number of Detained adults (n = 233) and Community adults (n = 233). The proportionality of the age groups was about the same for the Detained and Community adolescents, although 11-13 year olds made up a slightly smaller proportion of the Detained youths (16.3%) than of the Community youths (24.4%).

Males comprised 66.3% of the Detained sample and 56.8% of the Community sample, with approximately the same gender proportions in each age group within the Detained and the Community samples. African-American participants accounted for 40% of the sample, Hispanic 23%, non-Hispanic white 35%, and Asian 1%, with 1% of the participants not members of any of the above ethnic groups. These proportions are very similar to the ethnic proportions reported in a national survey of juvenile detention centers by the federal government a few years prior to data collection for the present study (43% African-American, 19% Hispanic, 35% non-Hispanic white: Snyder & Sickmund, 1995). **Table 1** shows that these proportions are apparent in each age group with the exception of the Detained 11-13 year old group, in which African-American youths were slightly over-represented and non-Hispanic white youths slightly under-represented. **Table 2** provides further information on gender-by-ethnicity proportions within age groups for the Detained and Community samples.

Finally, *Table 1* shows that the participants were classified predominantly (60%) in the two lowest (poorest, least educated) socioeconomic categories (IV and V in the Hollinghead system) and only 10% in the two highest categories (I and II). These proportions were similar for each age group. They were also similar across Detained samples and across Community samples, but there were slightly greater proportions in categories IV and V among Detained youth than among Community youth.

### *Sites*

Data were collected in four geographic areas of the U.S. by four research teams:

- Los Angeles County, CA: University of California at Los Angeles (n = 404; 29% of total sample)
- Alachua County, FL: University of Florida (n = 223; 16%)
- Philadelphia, PA: Temple University (n = 390; 28%)
- Albemarle, Alexandria and Fairfax Counties, VA: University of Virginia (n = 376; 27%)

Across these four geographic sites, participants were obtained from a total of 11 juvenile detention centers and 8 adult jails, as well as from the communities that were served by those facilities. *Table 3* shows the relative contributions of each of the four sites to the overall sample, as well as proportional contributions by gender, ethnicity and age-by-Detained/Community groups. The Florida site contributed a smaller proportion to the overall sample than the other three sites (due to a later starting data for data collection) and, therefore, proportionately less to each of the demographic groups. The proportional contributions of the other three sites were very similar with respect to all demographic variables, with the exception of ethnicity. The California site contributed a greater proportion of Hispanic cases, while the

Virginia site contributed a greater proportion of non-Hispanic white cases. In addition, the Florida site contributed somewhat disproportionately to the 11-13 year old Community subsample.

### ***Detained Samples: Legal Variables***

All Detained participants were residing in secure detention or jail facilities during a pretrial phase of their legal cases. **Table 4** shows that a variety of legal charges were represented among the Detained samples, with the adult Detained group having a somewhat greater proportion of drug charges and the adolescent Detained group a somewhat greater proportion of charges for offenses against persons. **Table 4** also shows Detained participants' responses to two questions (see "Measures") designed as indices of their prior involvement with the justice system. Except for the youngest group, about half of the Detained participants had been both locked up before and adjudicated delinquent or guilty on earlier charges.

### ***Samples: Other Independent Variables***

Other independent variables included intelligence and mental status. Ordinarily these would be reported at the beginning of a paper's Results section, but they are offered here as participant description for archival purposes. **Table 5** shows IQ scores (*Wechsler Abbreviated Scale of Intelligence*) and mental/emotional status (*Massachusetts Youth Screening Instrument-Second Version: MAYSI-2*), for age groups within the Detained and Community samples.

The IQ mean for Detained youths was significantly lower than for Community youths [DY mean = 85.58 (s.d. = 12.31), CY mean = 96.41 (s.d. = 15.22);  $t = -11.88$ ,  $df = 900$ ,  $p < .01$ ] and was significantly lower for Detained adults than for Community adults [DA mean = 87.65 (s.d. = 14.05), CA mean = 99.59 (s.d. = 16.25);  $t = -8.48$ ;  $df = 454$ ;  $p < .01$ ]. There were small

but statistically significant differences in IQ scores between Detained youths and Detained adults [DY mean = 85.58 (s.d. = 12.31), DA mean = 87.65 (s.d. = 14.05);  $t = 1.99$ ,  $df = 684$ ,  $p < .05$ ] and between Community youths and Community adults [CY mean = 96.41 (s.d. = 15.22), CA mean = 99.59 (s.d. = 16.25);  $t = 2.56$ ,  $df = 705$ ,  $p < .02$ ].

MAYSI-2 scores were generally higher for Detained than for Community participants. Tests of significance of difference in means between all Detained youths and all Community youths for each MAYSI-2 scale indicated that Detained youths scored significantly higher than Community youths on Alcohol/Drug Use ( $t=18.48$ ), Angry/Irritable ( $t=10.51$ ), Depressed/Anxious ( $t=10.48$ ), Somatic Complaints ( $t=4.83$ ), Suicide Ideation ( $t=6.5$ ) and Thought Disturbance ( $t=5.93$ ). Similarly, significantly higher mean scores were found for Detained adults than for Community adults on Alcohol/Drug Use ( $t=13.62$ ), Angry/Irritable ( $t=5.21$ ), Depressed/Anxious ( $t=6.59$ ), Somatic Complaints ( $t=3.19$ ), Suicide Ideation ( $t=6.7$ ) and Thought Disturbance ( $t=5.01$ ).

### *Measures for Independent Variables*

#### *Demographic Variables*

The study obtained data on all participants regarding age, gender, ethnicity and socioeconomic status.

Ethnicity and race were determined by participant self-identification in response to an open-ended query ("What race do you consider yourself to be?") followed by a more specific query ("Do you consider yourself to be Hispanic or Latino?").

SES was determined on the basis of participant report of the education and occupation of the family or guardian with whom the participant most recently resided, with values assigned

according to the Hollingshead system (1975) (providing five hierarchical classes). Subjects who reported that they lived independently or with roommates, or that they were homeless, were asked to report the education and occupation of their family of origin. If the participant knew only the guardian's education, the mean occupation value for guardians with that education was substituted for missing data. If the participant knew only the guardian's occupation, the mean education value for guardians in that occupation was substituted. More than 90% of participants were able to provide adequate information for their guardians. Participants who could not supply any information for their most recent guardian were assigned the SES value of the individual for whom they were able to provide the most education or occupation information (e.g., biological parent, adoptive parent or step-parent). Only 17 subjects were not able to provide any education or occupation information about their guardians, parents or other relatives.

### ***Justice System Familiarity***

For the Detained sample, participants were asked whether they had (a) ever before been “found guilty” of a delinquency or crime and (b) ever before been “locked up” in a detention center or jail. These two questions were used to create an index of past exposure to justice settings, coding “no” to both questions as “0,” “yes” to only one of the questions as “1,” and “yes” to both questions as “2.”

### ***Wechsler Abbreviated Scale of Intelligence (WASI)***

The WASI (Psychological Corporation, 1999) is a brief scale of intelligence that estimates full scale IQ using two or four subtests and has been normed for ages 6 through 89. This study used two-subtest format of Vocabulary and Matrix Reasoning which requires about

15 minutes to administer. The WASI two-subtest format FSIQ correlates .81 and .87 respectively with the Wechsler Intelligence Scale for Children-III and Wechsler Adult Intelligence Scale-III full scale IQ's.

### ***Massachusetts Youth Screening Instrument-Second Version (MAYSI-2)***

The MAYSI-2 (Grisso & Barnum, 2000; Grisso, Barnum, Fletcher, Cauffman & Peuschold, 2001) is a 52-item yes/no self-report instrument on which individuals are asked to report whether the various thoughts, feelings or behaviors described in the items have been true for them “in the past few months.” The items contribute to six scales that represent mental or emotional conditions that may warrant clinical attention (*Alcohol/Drug Use, Angry-Irritable, Depressed-Anxious, Somatic Complaints, Suicide Ideation, Thought Disturbance*). The instrument was designed for use in screening youths entering juvenile detention facilities, and research for its development indicated satisfactory internal consistency, reliability, and construct validity (when compared to other measures of adolescent psychopathology). It was administered in the present study to adults as well, the only modifications being replacement of the word “work” for “school” in two of the items.

Alpha coefficients (internal consistency of the scales) ranged from .61 to .86 in Massachusetts and California samples of adolescents in the juvenile justice system (Grisso et al, 2001). The scales correlated highly with comparable scales in more comprehensive measures of mental and emotional status (*Millon Adolescent Clinical Inventory and Child Behavior Checklist-Youth Self Report*: Grisso et al., 2001).

**Table 6** shows alpha coefficients for the six scales for Detained and Community youth and adults in the present study. These are comparable to the coefficients obtained in earlier

MAYSI-2 reports, and they are relatively similar for youths and for adults. A notable exception was *Thought Disturbance*, which did not achieve acceptable alpha coefficients for any of the subgroups (ranging from .53 to .13, average .408 across the eight subgroups).

Performance on the MAYSI-2 scales is expressed in several ways in the results of this study: (a) mean scores, (b) the percentage of participants in groups who exceed cut-off scores for “clinical significance,” and (c) the percentage of participants in groups who exceed cut-off scores representing the top 10% of scorers in the original Massachusetts norming study. Cut-off scores for “clinical significance” were determined in earlier studies in which youths had also taken the *Millon Adolescent Clinical Inventory* and *Child Behavior Checklist-Youth Self Report*, both of which have established cut-off scores for clinical levels of mental or emotional disturbance on scales within those instruments. The MAYSI-2 validation study (Grisso et al., 2001) used participants’ scores above or below those cut-offs to determine the most predictive cut-offs (called the “Caution cut-off score”) for the MAYSI-2 scales that were related conceptually to the comparison scales in the *Millon* and the *Youth Self Report*. Scores above the clinical significance (“Caution”) cut-off, therefore, may be interpreted as suggesting a high likelihood that those participants would score in the “clinically-significant range” on other, more comprehensive measures of adolescent psychopathology. This assumption, however, has not been tested with adults.

### ***Measures for Dependent Variables***

Two measures were used to assess abilities and dispositions related conceptually to youths’ and adults’ capacities as trial defendants. The first was associated more formally with the concept of competence to stand trial. The second was developed for the present study to

examine more broadly questions of psychosocial maturity and individuals' decision making in the adjudicative process.

### ***MacArthur Competence Assessment Tool-Criminal Adjudication (MacCAT-CA)***

The MacCAT-CA (Poythress, Nicholson, Otto, Edens, Bonnie, Monahan, & Hoge, 1999) was developed to assess abilities associated with "adjudicative competence," focusing on factors that may be involved in legal determinations of defendants' "competence to proceed (assist counsel)" and their "decisional competence" (Bonnie, 1992). The MacCAT-CA was the product of a substantial research project designed to produce an assessment tool that could be used to obtain information about defendants' abilities associated with their participation as defendants in their trials ("competence to stand trial") (Hoge, Bonnie, Poythress, Monahan, Feucht-Haviar, & Eisenberg, 1997; Otto, Poythress, Edens, Nicholson, Monahan, Bonnie, Hoge, & Eisenberg, 1998). The scales and items were developed to be consistent with conventional psycholegal interpretations of the legal criteria for competence to stand trial. The instrument was normed on large, national samples of pretrial adult defendants, including subsamples of defendants found incompetent to stand trial, defendants not found incompetent but being treated for mental illness, and defendants without evidence of incompetence or mental disorder. Norms and scoring criteria are provided in the manual (Poythress et al., 1999). At the time of the present study, there were no publications reporting use of the MacCAT-CA with adolescents. The present study made no modifications in the instrument in preparation for its administration to adolescents.

The MacCAT-CA uses a structured and highly standardized interview that consists of 22 items organized in three parts called *Understanding*, *Reasoning*, and *Appreciation*. The 8

Understanding items and 8 Reasoning items are based on responses to questions about a brief vignette, introduced at the beginning of the interview, regarding two men who get in a fight in a bar while playing pool. One of the men is arrested and faces charges and a trial process.

*Understanding.* In the hypothetical context presented in the vignette, the Understanding items ask questions of the examinee to determine the examinee's comprehension of 8 things:

- roles of defense and prosecution attorneys
- the elements of an offense with which the defendant could be charged
- elements of a lesser included offense
- role of the judge
- role of the jury
- the consequences of conviction
- consequences of pleading guilty
- rights waived when one pleads guilty.

For 6 of these items, an examinee's failure to give an adequate response results in the examiner employing a brief, standardized "teaching" process regarding the element that is being tested, with a subsequent repetition of the question to determine whether the defendant can provide an adequate response after the "teaching" has occurred. For the 2 items about nature of the offense, the items begin with a "teaching" disclosure without a "pre-test" of the defendant's knowledge.

The examinee's responses are scored using objective scoring criteria in the manual, classifying them as 2, 1, or 0 (adequate, questionable, inadequate). The Understanding score for each of the 6 items that have a "pre-test" and a potential "teaching" section is the examinee's

score on the “pre-test” if it was adequate (2 pt.) or, if teaching was needed, the score obtained after the teaching process. MacCAT-CA Understanding scores range from 0 to 16.

In the present study, two scores based on the Understanding items were calculated: an “Unaided Understanding” score (UU), and the standard MacCAT-CA Understanding score as described above. The UU score was based on participants’ original responses, prior to “teaching,” on the 6 items that have a potential teaching process (items 1, 4, 5, 6, 7, 8). They are intended to represent the participant’s degree of understanding of content assessed by the MacCAT-CA Understanding process without assistance that might augment their comprehension.

**Reasoning.** The 8 Reasoning items are of two types. Five of the items assess the individual’s recognition of the relevance of factual information for a defense. Each item presents two “facts” about the incident and asks the examinee which of the two would be more important for the defendant to tell his lawyer. The other 3 items were designed to assess a participant’s ability to process information when making a decision in a legal context. The items involve a description of two pleading choices: pleading guilty under the conditions of a plea agreement with the prosecutor, and pleading not guilty and going to trial.

The items are scored according to evidence that the examinee (a) seeks more information before deciding, (b) offers both an advantage and a disadvantage for the chosen option, and (c) has engaged in a comparison of the chosen option to the rejected option. (Scoring does not evaluate the specific choice that the examinee has made; scores focus entirely on the three factors above that reflect the *process* by which the choice was made.) Responses are scored (2, 1, 0) according to objective criteria provided in the manual, allowing MacCAT-CA Reasoning scores to range from 0 to 16.

**Appreciation.** The 6 Appreciation items were designed to assess whether defendants' perceptions of their own legal situations manifested illogical or implausible beliefs suggestive of the influence of mental disorder (e.g., delusional thinking). The items are not based on the hypothetical case used in Understanding and Reasoning, but on the examinee's own legal situation. The examinee is asked whether, "compared to other people who are in trouble with the law," the examinee thinks that he or she is "more likely, less likely, or just as likely" to:

- be treated fairly in the legal process
- be assisted by defense counsel
- fully disclose case information to his or her defense attorney
- be found guilty
- get the same punishment as others if found guilty
- plead guilty

Examinees are then asked for an explanation for their choice, and queries are made to determine whether the examinee's explanation contains evidence of "unrealistic or idiosyncratic beliefs that defendants have about themselves or their situations" such that they are "clearly implausible and colored by symptoms of mental illness (e.g., delusions)" (Poythress et al., 1999, p. 13). Responses are scored not on the basis of the choice, but according to the explanation that is provided. Items are scored according to objective criteria provided in the manual, including 2 point credit for responses that contain no implausible beliefs, 1 point for responses that are ambiguous regarding implausible beliefs, and 0 credit for responses that clearly contain implausible beliefs or where the interviewee provides an answer but does not know his or her reasons for it. MacCAT-CA Appreciation scores range from 0 to 12.

The MacCAT-CA as used in this study, therefore, provided four scores: Unaided Understanding, Understanding, Reasoning, and Appreciation. Consistent with the MacCAT-CA standardized procedure, no summary score or “total MacCAT-CA score” was calculated.

Inter-rater reliability for the MacCAT-CA was assessed at two points in the present project, employing all scorers at all four of the study sites (as described later in “Procedure”). At each time point, 25 research assistants scored 40 protocols. Initial inter-rater reliability data were collected immediately after the research assistants were trained, using MacCAT-CA data collected from youths (see **Table 7**). Reliability data again were collected approximately six weeks prior to the conclusion of data collection, using data from youth and adult participants (see **Tables 8 and 9**). Because of the problem of high agreement and low kappa (see Feinstein & Cicchetti, 1990; Cicchetti & Feinstein, 1990), the observed and chance agreement values are presented.

For youths’ protocols, intra-class correlations for the three MacCAT-CA subscales were generally acceptable at both the beginning and end of the study, although they were much better at the end. The relatively low intra-class correlation for Appreciation at the beginning of the study was in the very acceptable range at the end of the study.

For adults’ protocols, intra-class correlations for Understanding and Reasoning were consistent with those obtained in the original study with the MacCAT-CA (Otto, et al., 1998). The intra-class correlation for Appreciation, however, was very low, as were kappa’s for the individual Appreciation items. This apparently was due to the fact that scores on Appreciation items for this sample of adults were highly truncated, with almost no adult participants producing 0-credit responses (almost all receiving 1 or 2 point credit).

The present study employed mean scale scores on the three MacCAT-CA scales, as well as a system of classifying MacCAT-CA scale scores into three hierarchical categories using cut-off scores provided in the MacCAT-CA manual. The cut-off scores identify persons with “minimal or no impairment,” “mild impairment,” or “clinically significant impairment.” The developers of the MacCAT-CA set these cut-off scores on the basis of criteria associated with statistical infrequency. The cut-off score for “clinically significant impairment” was set at the score equaling 1.5 standard deviations below the mean of the “presumed competent” samples in the original norming study. Performance above 1.0 standard deviation below the mean for those samples were considered to represent “minimal or no impairment. Scores between those two cut-offs were labeled “mild impairment.”

*A score in the “clinically significant impairment” range does not represent “incompetence to stand trial.” It merely indicates that the defendant performed at a level that is extremely low on the original norming study’s distribution of scores among adult defendants awaiting trial. Research with the MacCAT-CA, however, has indicated that scores below those cut-offs are made far more often by adult defendants who have been adjudicated incompetent to stand trial than by adult defendants who have not.*

### ***MacArthur Judgment Evaluation (MacJEN)***

This instrument, developed specifically for the present study, was intended to assess youths’ and adults’ decision making in the context of legal circumstances that face defendants. Its generally format was patterned after earlier instruments with similar intent (Woolard, Reppucci & Scott, 1996; Grisso, 1981). The instrument had two objectives: (a) to assess examinees’ choices in three legal decision contexts often facing defendants, and (b) to identify

and examine their explanations for those choices according to several dimensions of psychosocial maturity, as already described in the “Overview” of the present document. A manual for the instrument is in preparation.

Instrument development began by selecting three legal decision contexts often facing individuals in the delinquency/criminal process. These included:

- Decisions about one’s response to police officers’ requests to waive rights to silence and counsel in their pursuit of a statement from a suspect (Police Interrogation)
- Decisions about one’s respond to one’s defense attorney who requests information in preparation for a defense (Consulting Attorney)
- Decisions about one’s response to a “plea agreement” in which pleading guilty offers the prospect of a lesser penalty than the prospective outcome of pleading not guilty (Plea Agreement).

Brief vignettes were devised for each of these contexts, describing a third party faced with a relevant circumstance. For example, the first context (Police Interrogation) describes a suspect who has been taken into custody by police officers and is about to be questioned by them. The suspect is described as having been a lookout for others engaged in a crime. The police officers are asking the person to waive the right to silence. The examinee is asked to advise the suspect concerning what to do.

Then, after each context vignette, a set of structured interview questions were devised to lead the examinee through a process of:

- identifying ways the person in the vignette could respond to the situation
- eliciting examinees’ choice of a “best” and “worst” response among a set of options

- eliciting the examinee's explanation or reasons for believing that these choices are "best" and "worst"
- ranking the importance and impact of a set of consequences provided for the best and worst choices
- identifying what the examinee him/herself believes that he/she would do under the circumstances

The set of interview questions that cover the above procedures is almost identical for the Police Interrogation and Plea Agreement contexts, but is abbreviated for the Consulting Attorney context. Examinees' responses to the interview items contribute to several MacJEN variables as follows.

**Choices.** The first set of variables measures the examinee's choices of the "best" and "worst" option in each of the three decision contexts. These context-specific choice variables are not scored per se, but instead represent categorical decision outcomes. In the Police Interrogation vignette, respondents choose between (a) talking and admitting involvement, (b) denying involvement, and (c) remaining silent. Options for Consulting Attorney include (a) talking and admitting everything, (b) telling some of the story, (c) denying involvement, and (d) refusing to talk. In the Plea Agreement vignette, respondents simply choose between taking the plea or refusing the plea and going to trial. For each vignette, respondents are asked what the vignette character should do as well as their own choice if faced with a similar situation.

**Authority Compliance.** In each vignette, one decision choice represents compliance with authority: confessing to police, fully disclosing to the attorney, and accepting the prosecutor's plea agreement. An *authority compliance* score (AutComp) sums the number of compliant choices made across the three vignettes.

***Optimal Defense Orientation.*** In each vignette, one decision choice represents an optimal defense orientation: remaining silent during police interrogation, fully disclosing to the attorney, and accepting the prosecutor's plea agreement. An *optimal defense orientation* score (DefOrien) sums the number of defense oriented choices made across the three vignettes.

***Risk Appraisal.*** The risk appraisal variables are of three types. The first is *risk recognition*, or a person's ability to identify potential risks in a given situation. The second is *risk likelihood*, or a person's beliefs about the likelihood that possible negative consequences might occur. The third is *risk impact*, or the person's perceptions of how unpleasant the negative consequences would be if they did occur. The following variables are constructed from responses to the Police Interrogation and Plea Agreement vignettes. The Attorney Consultation vignette was abbreviated and did not include questions to assess these dimensions.

*Risk Recognition* is based on questions about the best and worst choices in each vignette. After the best choice is identified (e.g., talking and admitting to the police), participants are asked to identify all the good things and bad things that might happen as a result. The questions are repeated for the respondent's worst choice. Two scores are calculated across vignettes.

A "Risk Recognition 1" score (R-Rec1) represents the total number of risks identified across the best and worst choices in each vignette and then averaged across vignettes. A "Risk Recognition 2" score (R-Rec2) represents the percentage of a person's total consequences (good and bad) that were risks (bad). The R-Rec2 score for the two stories are averaged to produce a final R-Rec2 score between 0 and 100 percent. For both R-Rec1 and R-Rec2, higher scores represent greater recognition (absolute or percentage) of potential risks in legally relevant situations.

*Risk Likelihood* is based on questions about a standard set of potential consequences in each vignette. It is important to note that the measure does not indicate whether individuals are more or less accurate in their appraisals of the negative consequences. Rather, it is designed to allow relative comparisons of appraisals among groups. Risk Likelihood is operationalized through two variables: a "Risk Appreciation 1" score (R-App1) indicating the likelihood that possible negative consequences may occur and a "Risk Appreciation 2" score (R-App2) measuring how unpleasant the negative consequences would be if they did occur.

The R-App1 score is based on questions that ask respondents to rate whether a positive or negative outcome is more likely (e.g., police will let him go home or put him in detention; he will go to trial quickly or have to wait for his trial). Responses are scored on a four point scale anchored by the positive outcome definitely happening (1 pt) and the negative outcome definitely happening (4 pts.). Vignette-specific sum scores across four outcome pairs in the Police Interrogation vignette and five outcome pairs in the Plea Agreement vignette are averaged to create the R-App1 score. Scores may range from 4 to 18 where higher scores represent a greater likelihood that risks will occur.

The R-App2 score uses a four point scale to rate how unpleasant or "bad" it would be if each of the negative outcomes listed in R-App1 actually happened. Ranging from "okay" (1 pt.) to "very bad" (4 pts.), outcome ratings are summed within each vignette, and then averaged across vignettes to create a total score from 4 to 18. Higher scores represent appraisals of worse or more unpleasant impact.

**Future Orientation.** Future orientation represents a person's ability to identify potential long term consequences in a given situation. The following variables are constructed from responses to the Police Interrogation and Plea Agreement vignettes. The Attorney Consultation

vignette was abbreviated and did not include questions to assess these dimensions. Three scores are calculated across vignettes.

A "*Future Recognition 1*" score (F-Rec1) is based on questions about the best and worst choices in each vignette. For both choices, participants are asked to identify all the good things and bad things that might happen as a result. These responses were coded as reflecting short-range or long-range nature of their consequences. The F-Rec1 score represents the total number of long term consequences (good and bad) identified for both the best and worst choices in each vignette and then averaged across vignettes. A "Future Recognition 2" score (F-Rec2) represents the percentage of a person's total consequences (short term and long term) that were long term. The F-Rec2 score for the two stories are averaged to produce a final F-Rec2 score between 0 and 100 percent. For both F-Rec1 and F-Rec2, higher scores represent greater recognition (absolute or percentage) of potential long term consequences in legally relevant situations.

A "Future Recognition 3" score (F-Rec3) is based on a question in which participants are asked to provide the main reason why their best choice is better than their worst choice for each vignette. That main reason is coded as reflecting short term or long term consequences. The F-Rec3 score sums the number of long term main reasons across vignettes. Higher scores represent greater recognition of potential long term consequences.

***Resistance to Peer Influence.*** Resistance to peer influence compares the participants' own choice for themselves in each decision making vignette to their choice in a similar situation but under a condition of peer influence. After the entire MacJEN instrument is completed, participants are given a set of three vignettes in which peers recommended the opposite course of action (e.g., if a participant stated that they would admit to the police in the original MacJEN vignette, peers recommend remaining silent). The opposites are defined in the Police

Interrogation vignette as talking and admitting versus denying involvement or remaining silent; in the Consulting Attorney vignette as talking and admitting everything versus partial disclosure, denying involvement, or refusing to talk; and in the Plea Agreement vignette as refusing the plea versus taking the plea.

For each decision making vignette, resistance to peer influence is measured as a dichotomous variable (resisted influence and retained original choice versus influenced by peers and switched to peers' choice). A "Resistance to Peer Influence" score (RPI) sums the number of choices in which participants resisted peer influence across the three vignettes. Higher scores represent greater resistance to peers in legally relevant situations.

### *Measures of Psychosocial Maturity*

The main purposes of the study required comparisons between adolescents and adults on competence-related abilities—represented by our dependent measures—with regard to age, intelligence, and other concrete variables—represented by our independent measures.

However, a third type of measure used in this study represented variables that were neither strictly “independent” or “dependent,” but played both a descriptive role and a potentially explanatory role with regard to our main results. This included three measures representing concepts associated with “psychosocial maturity.”

Conceptualization of this study suggested that we should examine the hypothetical relation between psychosocial maturity and adolescent and adult defendants' decisions in legal contexts. We anticipated that adolescents' decisions about matters related to legal circumstances of defendants would be influenced by fundamental cognitive and social features of adolescent development. We expected that differences between adolescents' and adults'

performance on our dependent measures of decision making (and between various adolescent subsamples) would be associated with theoretically consistent differences in levels of maturity on dimensions of cognitive and social development.

As noted in the description of the MacJEN, we developed indices within the MacJEN to represent three dimensions of psychosocial maturity: (a) consideration of future consequences, (b) perception of potential negative consequences, and (c) resistance to peer influence. The MacJEN indices representing these concepts were expected to be useful in examining how these factors are potentially manifested in (and related to) youths' processing of decisions. But conceptually they might not serve as well for purposes of identifying youths' and adults' degrees of psychosocial maturity as a generalized construct (that is, outside the legal decision making process).

Therefore, we developed three measures that were intended to assess generalized dimensions of psychosocial maturity. As described below, the measures that were intended to represent them were:

- *Future Outlook Inventory*
- *Perceptions of Risk Scale*
- *Resistance to Peer Influence*

We expected to examine the relations of these measures to three other types of data in the study. *First*, they would gain construct validity if they were related to age, intelligence, and Detained/Community status in a manner consistent with developmental theory. *Second*, their comparison to the MacJEN measures of the same constructs would address questions of concurrent validity (anticipating at least a modest correlation between two measures of the same construct in somewhat different contexts). *Third*, the measures could be examined for their

relation to, and ability to account for, differences between the study's adolescents and adults in choices and reasoning on the MacJEN.

***Future Orientation: The "Future Outlook Inventory"***

For this study, we developed a measure of future orientation, focused on individuals' inclinations to recognize and consider future events or consequences in everyday life. A set of 14 items was drawn from the *Life Orientation Task* (Scheier & Carver, 1985), the *Zimbardo Time Perspective Scale* (Zimbardo, 1980); and the *Consideration of Future Consequences Scale* (Strathman, Gleicher, Boninger, & Edwards, 1994). Sample items included, "I usually think about the consequences before I do something" and "I would rather save my money for a rainy day than spend it now on something fun" where participants could respond to each item ranging from 1 ("Never True") to 4 ("Always True"). Initial pilot work with these items suggested that the instrument was reliable ( $\alpha = .71$ ) and correlated with self-reports of risk perception among high school and college students

An exploratory factor analysis indicated that the 14 items of the instrument loaded on two factors: one reflecting individuals' orientation to future (8 items) and one reflecting individuals' impulsivity and spontaneity (4 items). The remaining 2 items did not contribute to either factor. Confirmatory factor analyses indicated that the 8 items concerning individuals' orientation to future loaded on a single underlying factor in the detained and community, adolescent and adult samples. Accordingly, these 8 items were retained for The *Future Outlook Inventory* (FOI; Cauffman & Woolard). The FOI has a response scale for each item ranging from 1 ("Never True") to 4 ("Always True"), with scale scores ranging from 8 to 32. In the present study sample, internal consistency of the measure was  $\alpha = .66$ .

***Risk Perception: The “Perceptions of Risk Scale”***

Items for a risk perception measure called the *Perceptions of Risk Scale* were adapted (with permission) from an instrument developed by Benthin et al. (1993). “Risk perception” refers to the extent to which an individual recognizes and evaluates the risks inherent in activities that are potentially dangerous or harmful. The measure employed in this study presents the respondent with six activities:

- riding in a car with a drunk driver
- having unprotected sex
- smoking cigarettes
- stealing from a store
- drinking alcohol, and
- vandalizing property.

The procedure asks the respondent to indicate four things for each of these activities:

- how “scary” the activity is (affective component)
- how risky the activity is (likelihood component)
- how much the risks of the activity outweigh its benefits (comparative value component),  
and
- how serious the consequences of the activity would be if something “bad” happened as a result (salience component)

Each of these ratings is made on a 4-point scale by the examinee; evaluations of riskiness and the relative risk-benefit ratio are reverse-scaled and reverse-scored.

An exploratory factor analysis indicated that the four scales (affective, likelihood, comparative value, and salience) were part of one underlying factor, but that respondents’ ratings

were domain-specific, such that ratings of the two criminal activities (stealing and vandalizing) loaded on a different factor than ratings of the other four activities. In light of concerns that respondents' ratings of the criminal activities might be confounded with their membership in the detained versus community samples, we decided to drop the ratings of stealing and vandalizing and create a measure of risk perception based on the four evaluations (affective, likelihood, comparative value, and salience) of four activities (riding with an drunk driver, having unprotected sex, drinking, and smoking). A confirmatory factor analysis indicated that single underlying factor adequately fit the data for the detained and community, adolescent and adult samples. A single risk perception score was computed by averaging the 16 responses (four evaluation dimensions for four activities). In the study sample, the internal consistency of the PORS was  $\alpha=.86$ .

***Peer Influence: The “Resistance to Peer Influence” Scale***

Resistance to peer influence refers to an individual's capacity to stand up to peer pressure and behave as he or she wishes. Conventional measures of resistance to peer influence (e.g., Berndt, 1979; Steinberg & Silverberg, 1986) present the respondent with a series of hypothetical dilemmas, typically involving the possibility of engaging in some sort of antisocial activity (e.g., shoplifting, using illegal drugs, vandalizing property), and inform the respondent that although a group of friends wants the respondent to engage in the activity, he or she does not want to. The respondent is then asked whether he or she would or would not go along with the wishes of the peers. In the context of the present study, this approach was not suitable, for several reasons. First, the hypothetical items used in previous research were of questionable applicability to young adults. Second, because such a large proportion of the sample had histories of antisocial

activity, the format of existing measures asked respondents to agree to a premise (that they opposed the antisocial activity) that was contrary to fact. Finally, it was likely that, in a study in which half the sample knew they were recruited because of their prior antisocial behavior, respondents would feel a strong social desirability press to present themselves as resistant to pressure to engage in antisocial activity.

Accordingly, we developed a new measure, *Resistance to Peer Influence* (RPI), that was suitable for use with adults as well as adolescents, did not focus on antisocial behavior, and did not present resistance to influence as the obviously desirable response. A ten-item measure was developed using the item format developed for the *Self Perception Profile* (Harter, 1985). Each item contains two opposing statements, and the respondent is asked to designate which statement is more like him or her and indicate the degree of the item's applicability. For example, one item reads as follows: "Some people think it's more important to be an individual than to fit in with the crowd BUT other people think it is more important to fit in with the crowd than to stand out as an individual." The respondent is asked which statement is true about himself or herself, and whether it is "really true" or "sort of true." Each item is scored on a 4-point scale, with a score of 1 designating that the respondent has indicated that the statement reflecting the least resistance to peer pressure is "really true" and a score of 4 designating that the respondent has indicated that the statement reflecting the greatest resistance to peer pressure is "really true." Statements are written so that in some cases, the pressure-resistant statement is presented first and in others it is presented last.

Data collected from the study sample was subjected to confirmatory factor analysis, which indicated that a single underlying factor adequately fit the data for the detained and

community, adolescent and adult samples. In the study sample, the internal consistency of the RPI was  $\alpha = .74$ .

### *Procedure*

#### *Site Administration and Development*

A Working Group, functioning as an arm of the MacArthur Research Network on Adolescent Development and Juvenile Justice, oversaw the development and implementation of the study. Among its members were Network members as well as the Director of the study and the four site directors.

Each of the site directors was responsible for a site data collection team based at a university. Consistency of procedures across sites was the responsibility of a study Director and a Project Coordinator located at a university other than the data collection sites. Site directors, who were senior research faculty, were typically assisted by a senior research associate who shared responsibility for maintaining weekly or daily contact with the study coordinating site as well as providing direct and continuous supervision of the team's research assistants. A total of 25 graduate students or senior undergraduates in behavioral sciences participated as research assistants in this study, almost all of whom continued throughout the 18-month data collection period.

Each site was responsible for the development of organizational liaisons that were necessary in order to recruit participants for the study. Considerable pre-study effort was required to develop relationships with various juvenile detention centers and adult jails where data collection with Detained participants might be accomplished, as well as community organizations where the Community participants could be recruited. This complex process involved the development of agreements that would accommodate the daily needs and demands

of administration, staff and residents of the detention centers and jails, while also providing conditions that would assure the integrity of the data to be collected.

### ***Assurance of Ethical Research Practice***

The coordinating site, in consultation with the data collection sites, developed a prototype human subjects protocol that was first approved by the coordinating site's Human Subjects Committee. This protocol became the template for similar human subjects applications at the universities of each of the data collection sites. Similar applications were required by some of the agencies and facilities where data were to be collected. Not all sites' Human Subjects Committees arrived at exactly the same conditions for protection of the adolescent participants in the study, but the variations were not of a type that were likely to produce bias in samples across sites. (Of note, one university's Human Subjects Committee at a fifth site required procedures for youth assent that would have made the study impossible to conduct within the time and financial resources available for data collection, and that site had to be abandoned.) Youths in detention centers, as well as adults in jails, were identified uniformly as "vulnerable" populations. In general, the study was perceived by Human Subjects Committees as presenting low risk and low to mild benefit to participants, with substantial potential benefits to society and to persons like the participants.

Details of the conditions for the protection of human subjects will be made available in another document. Briefly, all participation was on a voluntary basis. Detained youths and adults received \$10 for participating in the study; community youths and adults received \$25. Informed and signed consent (assent for youths) was required for all adults and youths. Potential participants were informed of the purpose of the study, the nature of the procedures, the

time commitment, and the opportunity to withdraw if they volunteered. Confidentiality was assured with the exception of the researchers' obligation to report to staff or other authorities if the participant were to say or do things that suggested a significant risk of imminent self-harm, of being in danger of harm from others, or of harming others. The project obtained a federal certificate of confidentiality to assist in resisting potential requests by authorities (e.g., courts, attorneys) for disclosure of data on specific participants.

Human subjects considerations for adolescents required that parent consent be obtained for youths who were in the community (the Community youth sample), but not for youths in detention centers (the Detained youth sample). In compliance with Federal requirements for the protection of minors in residential facilities, independent "participant advocates" monitored the solicitation process in juvenile detention centers, assuring conditions of voluntary youth assent and vetoing specific youths' participation in circumstances in which the advocate believed that the youth might be under substantial stress at the time of the recruitment. Some facilities required the site teams to notify the parents of youths in detention centers concerning the intent to solicit the youth's voluntary participation, offering them an opportunity to object (in which case the youth was not approached by the research team). Note that in these instances, we did not presume that parents had given "passive consent" if they did not respond. (That is, parent consent—passive or active—was not required for detained youths at any of the sites, given the specific conditions of this study.)

### ***Team Training and Monitoring***

All data collection teams were trained extensively prior to beginning data collection, using several steps. First, site directors and research assistants from all teams met together at

one site for several days of basic instruction in the study's design, methods, measures, orientation regarding research ethics, and observation of mock administrations of the study's instruments. Second, research teams engaged in practice administrations among themselves under the tutelage of the site directors. Third, the project coordinator spent several days at each site, observing and providing feedback to all research assistants in their administration of the instruments and the entire research protocol to youths at the detention facilities where data collection would eventually occur. Fourth, all research assistants engaged in practice scoring exercises for the MacCAT-CA, MacJEN, and WASI. Fifth, as described earlier, all research assistants participated in the independent scoring of 40 MacCAT-CA protocols for purposes of determining interscorer reliability and identifying outlying scorers who might need further instruction.

### ***Recruitment***

***Detained Youths and Detained Adults.*** Research assistants for a given site visited the participating juvenile detention centers and adult jails weekly. Upon entry, they were assisted by staff to determine new detainees since the previous visit, and to determine whether any youths had been "screened out" by staff or youth advocates regarding potential participation. Any youths who had not been screened out were approached by research assistants with an explanation of the study, the procedure, and a request for assent to participate. As the study period progressed, approach to youths more selectively focused on specific ages or ethnicities that were needed to increase sample sizes within age/gender/race cells of the study's design.

***Community Youths and Community Adults.*** A variety of methods were used across the four geographic sites to obtain youth and adult volunteers from the community. For youths,

these included solicitation in schools, community youth programs, and girls' and boys' clubs. For adults, solicitation occurred in at community clubs, community agencies, shelters, and near community colleges. Solicitation was by posters, leaflets, and/or by workers at some of these programs. The solicitation targeted neighborhoods that would produce community samples with socioeconomic characteristics typical of those found among participants in the Detained samples. Solicitation and selection of community participants became more focused on specific gender and ethnicity as the study proceeded, using the demographic characteristics of the evolving Detained samples as a guide for producing Community samples with similar characteristics.

### *Administration of Protocol*

When a detained participant had consented/assented, the research assistant met with the participant in a room assigned by the detention center or jail, offering acceptable privacy and reduction of distractions. Demographic information was first obtained, followed by the MacCAT-CA, the WASI, the MacJEN, the MAYSI-2, and finally the Future Outlook Inventory, the Perceptions of Risk Scale, and Resistance to Peer Influence. This process typically required between 90 and 180 minutes.

Inasmuch as the MAYSI-2 had 5 items pertaining to potential self-harm, the project required that research assistants engage in further inquiry with participants who answered more than 2 out of 3 critical items pertaining to self-harm. Research assistants used a specific set of interview questions to probe the seriousness and recency of the participant's intent. Responses to the items were used to determine whether the research assistant would be required to report the results of those items to detention or jail staff, or parents in the case of Community youths. This decision was made by the site project manager after reviewing the interview responses with

the research assistance, and where necessary, in consultation with a site clinician and/or the project director at the coordinating site for this study. Inquiry was needed in 8% (n = 75) of youth cases and 4% (n = 21) of adult cases. Research assistants reported to staff or parents that there was a risk of imminent self-harm in 24% (n = 18) of the inquiry cases (2% of total sample) involving youths, and in 24% (n = 5) of the inquiry cases (1% of total sample) involving adults. In addition, less than 1% (n = 4) of Community Adults and less than 1% (n = 1) of Community Youths were given a list of phone numbers for local community mental health facilities and hotline services. Documentation of the inquiry, the site manager's decision, and supporting reasons were sent to the central coordinating site for the study for review and archiving.

### ***Monitoring for Quality Control***

The coordinating site (Univ. of Massachusetts Medical School) developed a standard set of research materials and protocols, conducted training at the data collection sites (as described earlier), continuously monitored the data collection process at the four data collection sites, and developed and managed the study's data base. Site managers were in contact with the project director at the coordinating site several times a week early in data collection and weekly thereafter, addressing issues of recruitment, data collection, and scoring. Research assistants at each site scored their own protocols, and scored copies were sent to the coordinating site for quality checks and data entry.

**Table 1: Sample Demographics**

	Detained					Community				
	Youth Age Groups			Youth	Adults	Youth Age Groups			Youth	Adults
	11-13	14-15	16-17	11-17	18-24	11-13	14-15	16-17	11-17	18-24
Participants (n)	74	186	193	453	233	116	159	199	474	233
Male (% of age group)	74	62	62	64	71	52	60	57	57	57
Ethnicity (% of age group)										
African-American	56	32	38	39	43	41	52	33	41	37
Hispanic	21	28	25	26	25	20	20	21	20	24
Non-Hispanic White	21	35	35	32	32	36	28	44	37	37
Asian and Other	2	5	2	3	0	3	0	2	2	2
Socioeconomic Status (% of age group)										
I-II	8	7	11	9	7	15	13	15	14	9
III	12	16	18	16	16	23	26	24	24	18
IV-V	80	77	71	75	77	62	61	61	62	73

**Table 2: Number of Participants in Gender/Ethnicity by Age Groups  
(Percent of Age Group in Parentheses)**

	Detained				Community			
	11-13	14-15	16-17	18-24	11-13	14-15	16-17	18-24
<b>Males</b>								
African-American	27 (37)	40 (22)	47 (24)	70 (30)	24 (21)	49 (31)	32 (16)	47 (20)
Hispanic	12 (16)	30 (16)	25 (13)	47 (20)	15 (13)	16 (10)	22 (11)	32 (14)
Non-Hispanic White	13 (18)	41 (22)	46 (24)	48 (21)	20 (17)	30 (19)	55 (28)	50 (21)
Asian and Other	2 (3)	4 (2)	2 (1)	0	1 (1)	1 (1)	4 (2)	3 (1)
<b>Females</b>								
African-American	14 (19)	20 (11)	27 (14)	29 (12)	23 (20)	33 (21)	33 (17)	39 (17)
Hispanic	3 (4)	23 (12)	23 (12)	12 (5)	8 (7)	15 (9)	19 (10)	24 (10)
Non-Hispanic White	2 (3)	24 (13)	21 (11)	27 (12)	22 (19)	15 (9)	33 (17)	37 (16)
Asian and Other	0	4 (2)	1 (1)	0	3 (2)	0	0	1 (1)

**Table 3: For Each Ethnic Category, Proportion of Samples Contributed by Each Site**

	Detained Subjects		Community Subjects	
	11-17 n (%)	18+ n (%)	11-17 n (%)	18+ n (%)
<b>African-American</b>				
California	55 (31.4)	53 (27.3)	7 (7.1)	29 (33.7)
Florida	21 (12)	35 (18)	30 (30.3)	7 (8.1)
Pennsylvania	56 (32)	55 (28)	26 (26.3)	25 (29.1)
Virginia	43 (24.6)	51 (26.3)	36 (36.4)	25 (29.1)
<b>Hispanic</b>				
California	62 (53.4)	62 (65.3)	47 (79.7)	35 (62.5)
Florida	7 (6)	5 (5.3)	0	4 (7.1)
Pennsylvania	29 (25)	25 (26.3)	10 (16.9)	16 (28.6)
Virginia	18 (15.5)	3 (3.2)	2 (3.4)	1 (1.8)
<b>Non-Hispanic White</b>				
California	15 (10.2)	13 (7.4)	12 (16)	8 (9.2)
Florida	23 (15.6)	45 (25.7)	14 (18.7)	18 (20.7)
Pennsylvania	39 (26.5)	55 (31.4)	25 (33.3)	24 (27.6)
Virginia	70 (47.6)	62 (35.4)	24 (32)	37 (42.5)

**Table 4: Legal Circumstances of Detained Participants  
(Percent within Age Groups)**

	11-13	14-15	16-17	18-24
<b>Current Charges</b>				
Property offenses	34	48	42	29
Person offenses	51	33	37	22
Drug charges	10	8	12	32
Probation violation/Other	5	11	9	17
<b>Involvement with Justice System</b>				
Reported “locked up” before	33	52	57	58
Reported “pled or found guilty” before	41	66	66	67
Percent “yes” on both of above	32	50	55	56

**Table 5: Psychological Measures by Age Groups**

	Detained				Community			
	11-13	14-15	16-17	18-24	11-13	14-15	16-17	18-24
<b>Wechsler Abbreviated Scale of Intelligence (IQ)</b>								
Mean	84.8	84.3	87.1	87.6	97.3	94.7	97.2	99.6
Standard deviation	13.3	12.4	11.7	14.0	16.2	14.6	15.1	16.2
Percent IQ 60-74	20	24	14	17	11	8	4	3
<b>Massachusetts Youth Screening Instrument-2</b>								
<i>Scale Means</i>								
Alcohol/Drug Use	2.47	2.97	3.55	3.77	0	0.56	0.78	1.09
Angry-Irritable	3.86	4.74	4.41	4.16	2.39	3.03	2.71	2.90
Depressed-Anxious	2.52	3.33	2.76	3.19	1.45	1.72	1.51	1.91
Somatic Complaints	2.63	3.04	3.00	3.05	2.17	2.45	2.46	2.48
Suicide Ideation	0.81	0.94	0.66	0.56	0.26	0.36	0.26	0.30
Thought Disturbance	1.00	0.88	0.79	0.90	0.50	0.55	0.45	0.47
<i>Percent over clinical significance criterion (MAYSI-2 "Caution" Cut-off)</i>								
Alcohol/Drug Use	24	43	52	44	0	6	9	11
Angry-Irritable	40	58	48	45	18	32	23	27
Depressed-Anxious	43	55	47	55	24	28	23	33
Somatic Complaints	54	56	60	63	41	48	49	49
Suicide Ideation	18	24	17	14	7	11	7	8
Thought Disturbance	54	52	54	56	35	37	29	34

**Table 6: MAYSI-2 Scales: Alpha Coefficients by Sample and Gender**

Sample and Gender	MAYSI-2 SCALES					
	ADU	Ang/Irr	Dep/Anx	Somatic	Suicide	Tht Dist
Detained Youths						
Females	.8401	.7284	.6839	.6764	.7566	.3674
Males	.8497	.7702	.7202	.6808	.7586	.5358
Community Youths						
Females	.7951	.6652	.6873	.6796	.7750	.5120
Males	.8257	.7315	.5784	.5160	.6360	.4246
Detained Adults						
Females	.7658	.7611	.6814	.7265	.7110	.3131
Males	.8216	.7566	.7183	.7252	.6777	.4887
Community Adults						
Females	.7882	.7456	.6166	.7228	.6993	.5357
Males	.7777	.7723	.6544	.6347	.5242	.1362

**Table 7: Initial Inter-Rater Reliability Coefficients (Intra-Class and Kappa),  
For MacCAT-CA Scales and Items**

Scale Items	Intra-Class Correlation	Kappa	P(Observed Agreement)	P(Chance Agreement)
Understanding	.6362			
Item 1		.7097	.9623	.8700
Item 2		.6327	.7930	.4364
Item 3		.5196	.7224	.4222
Item 4		.7825	.8977	.5297
Item 5		.5788	.7605	.4315
Item 6		.5512	.9254	.8337
Item 7		.8508	.9672	.7806
Item 8		.9040	.9447	.4239
Reasoning	.6044			
Item 9		.6724	.8807	.6350
Item 10		.6590	.8988	.7031
Item 11		.6208	.8096	.4981
Item 12		.6864	.8198	.4254
Item 13		.8413	.9354	.5927
Item 14		.6663	.7944	.3839
Item 15		.4717	.8819	.7764
Item 16		.3049	.5961	.4190
Appreciation	.4115			
Item 17		.3130	.8031	.7134
Item 18		.1345	.8345	.8088
Item 19		.2270	.7464	.6720
Item 20		.2069	.7965	.7434
Item 21		.4767	.8026	.6228
Item 22		.0181	.8435	.8406

**Table 8: Final (End of Study) Inter-Rater Reliability Coefficients (Intra-Class and Kappa), for MacCAT-CA Scales and Items (Youth Protocols Only)**

Scale Items	Intra-Class Correlation	Kappa	P(Observed Agreement)	P(Chance Agreement)
Understanding	.9108			
Item 1		.8211	.9192	.5484
Item 2		.5690	.7577	.4377
Item 3		.6362	.7579	.3346
Item 4		.7567	.9343	.7301
Item 5		.6195	.7609	.3715
Item 6		.7535	.8815	.5192
Item 7		.7391	.8943	.5948
Item 8		.7146	.8183	.3633
Reasoning	.8078			
Item 9		.7285	.8843	.5743
Item 10		.6705	.8618	.5806
Item 11		.4714	.7114	.4541
Item 12		.6419	.7826	.3928
Item 13		.8046	.9037	.5071
Item 14		.5668	.7072	.3241
Item 15		.6823	.8860	.6367
Item 16		.3267	.5794	.3754
Appreciation	.8606			
Item 17		.6585	.8846	.6622
Item 18		.5438	.7847	.5280
Item 19		.2307	.6648	.5642
Item 20		.4629	.806	.6388
Item 21		.6302	.8605	.6228
Item 22		.3530	.7105	.5526

**Table 9: Final Inter-Rater Reliabilities: Adult Protocols Only**

Scale	Intra-Class Correlation	Kappa	P(Observed Agreement)	P(Chance Agreement)
Understanding	.8843			
Item 1		-.003	.9916927	.9916667
Item 2		.7076	.8274	.4097
Item 3		.6610	.782	.3569
Item 4		.8471	.9739	.8293
Item 5		.6342	.7928	.4334
Item 6		.9272	.9791	.714
Item 7		.6618	.9602	.8823
Item 8		.8033	.8925	.4537
Reasoning	.7076			
Item 9		.8222	.9314	.6139
Item 10		.4148	.8436	.7328
Item 11		.3201	.7182	.5858
Item 12		.5237	.7493	.4738
Item 13		.7488	.8535	.4168
Item 14		.5240	.7082	.387
Item 15		.7192	.9228	.7251
Item 16		.2552	.5855	.4435
Appreciation	.1762			
Item 17		.1260	.7772	.7451
Item 18		.0592	.784	.7704
Item 19		.1018	.643	.6025
Item 20		.0836	.727	.7020
Item 21		.3044	.8241	.7471
Item 22		.0777	.7162	.6923

### *References*

- Benthin, A., Slovic, P., & Severson, H. (1993). A psychometric study of adolescent risk perception. *Journal of Adolescence, 16*, 153-168.
- Berndt, T. (1979). Developmental changes in conformity to peers and parents. *Developmental Psychology, 15*, 608-616.
- Bonnie, R. (1992). The competence of criminal defendants: A theoretical reformulation. *Behavioral Sciences and the Law, 10*, 291-316.
- Cicchetti, D., & Feinstein, R. (1990). High agreement but low kappa: II. Resolving the paradoxes. *Journal of Clinical Epidemiology, 43*, 551-558.
- Feinstein, A., & Cicchetti, D. (1990). High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology, 43*, 543-549.
- Grisso, T. (1981). *Juveniles' waiver of rights: Legal and psychological competence*. New York: Plenum.
- Grisso, T., & Barnum, R. (2000). *Manual for the Massachusetts Youth Screening Instrument-Second Version (MAYSI-2)*. Worcester, MA: University of Massachusetts Medical School.
- Grisso, T., Barnum, R., Fletcher, K., Cauffman, E., & Peuschold, D. (2001). Massachusetts Youth Screening Instrument for mental health needs of juvenile justice youths. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 541-548.
- Harter, S. (1985). *The self-perception profile for adolescents*. Unpublished manuscript, Department of Psychology, University of Denver.
- Hoge, S., Poythress, N., Bonnie, R., Eisenberg, M., Monahan, J., Feucht-Haviar, T., & Oberlander, L. (1996). Mentally ill and non-mentally ill defendants' abilities to understand information relevant to adjudication: A preliminary study. *Bulletin of the American Academy of Psychiatry and the Law, 24*, 187-197.

Hollingshead, A. (1975). *Four factor index of social status*. Unpublished manuscript.

Otto, R., Poythress, N., Edens, J., Nicholson, R., Monahan, J., Bonnie, R., Hoge, S., & Eisenberg, M. (1998). Psychometric properties of the MacArthur Competence Assessment Tool-Criminal Adjudication. *Psychological Assessment, 10*, 435-443.

Poythress, N., Nicholson, R., Otto, R., Edens, J., Bonnie, R., Monahan, J., & Hoge, S. (1999). *The MacArthur Competence Assessment Tool-Criminal Adjudication*. Odessa, FL: Psychological Assessment Resources.

Psychological Corporation (1999). *Wechsler Abbreviated Scale of Intelligence Manual*. San Antonio, TX: Psychological Corporation.

Scheier, M., & Carver, C. (1985). Optimism, coping and health: Assessment and implications of generalized outcome expectancies. *Health Psychology, 4*, 219-247.

Snyder, H., & Sickmund, M. (1995). *Juvenile offenders and victims: A national report*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Strathman, A., Gleicher, F., Boninger, D., & Edwards, C. (1994). The consideration of future consequences: Weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology, 66*, 742-752.

Steinberg, L., & Silverberg, S. (1986). The vicissitudes of autonomy in early adolescence. *Child Development, 57*, 841-851.

Woolard, J, Reppucci, N., & Scott, E. (1996). *Judgment Assessment Tool - Adolescents*. Charlottesville, VA: University of Virginia (unpublished manual).

Zimbardo, P. (1990). *The Stanford Time Perspective Inventory*. Stanford, CA: Stanford University.