Law and Human Behavior, Vol. 19, No. 2, 1995

The MacArthur Treatment Competence Study. III

Abilities of Patients to Consent to Psychiatric and Medical Treatments*

Thomas Grisso and Paul S. Appelbaum

Three instruments assessing abilities related to legal standards for competence to consent to treatment were administered to 6 groups: patients recently hospitalized for schizophrenia, major depression, and ischemic heart disease, as well as three groups of non-ill persons in the community who were matched with the hospitalized patients on age, gender, race, and socioeconomic status. Significant impairments in decisional abilities were found for only a minority of persons in all groups. Both the schizophrenia and depression groups manifested poorer understanding of treatment disclosures, poorer reasoning in decision making regarding treatment, and a greater likelihood of failing to appreciate their illness or the potential benefits of treatment. Deficits were more pronounced, however, among patients with schizophrenia. Implications are discussed for policy designed to protect the rights and welfare of patients with mental illness who are at risk of incompetent refusal or consent when making treatment decisions.

This article reports the method and results of the MacArthur Treatment Competence Study of mentally ill and medically ill patients' abilities to make decisions about their treatment. The purpose of this study was to produce reliable data

* This research was supported by the John D. and Catherine T. MacArthur Foundation Research Network on Mental Health and the Law. Other members of the Network are acknowledged for their conceptual guidance during the process of the study: Shirley Abrahamson, Richard Bonnie, Pamela Hyde, John Monahan, Stephen Morse, Ed Mulvey, Loren Roth, Paul Slovic, Henry Steadman, and David Wexler, as well as Steven K. Hoge. Acknowledged for their roles in data collection and statistical analysis are Kenneth Fletcher, Deidre Klassen and Cynthia Wickless. Requests for manuals for the MacArthur Treatment Competence Research Instruments, and other correspondence, may be addressed to either author at the Department of Psychiatry, University of Massachusetts Medical School, Worcester, MA 01655.

0147-7307/95/0400-0149$07.50/0 © 1995 Plenum Publishing Corporation

offering an empirical perspective for law, policy, and clinical decisions regarding patients' competence or incompetence to make treatment decisions. In a companion article (Appelbaum & Grisso, 1995), we described legal, historical, and conceptual issues of informed consent that have given rise to the need for empirical information on patients' decisional abilities. That article also reviewed relevant empirical research related to those issues and provided the rationale and design for the study to be described here. A second companion article (Grisso, Appelbaum, Mulvey, & Fletcher, 1995) described the development, reliability, and internal validity of three measures (the MacArthur Treatment Competence Research Instruments) to be used in research on patients' abilities related to four legal standards for determinations of competence to consent to treatment. Those measures were employed in the study reported here.

Specifically, this study had three major objectives. *First,* we wished to determine the extent to which mentally ill patients do or do not manifest deficits in legally relevant abilities to make treatment decisions, compared to the abilities of nonmentally ill patients and persons in the community with no illnesses. *Second,* if differences were found among these groups in the relevant abilities, we wished to identify any reliable demographic and clinical characteristics that would assist in describing those patients who might be at greater risk of incapacities for making treatment decisions (providing competent informed consent). *Third,* the study was designed to examine patterns of deficits in abilities related to various legal standards for competence.

METHODS

Participants

Data were collected from three groups of hospitalized patients and an additional three groups of persons who were not hospitalized and were residing in communities from which the hospitalized participants were obtained.

Hospitalized Groups

Two of the hospitalized groups were obtained from a mix of public and university hospital psychiatric units where they had been admitted for treatment of mental disorders. The *schizophrenia* group (n = 75) included patients whose admitting ("chart") diagnosis was schizophrenia or schizoaffective disorder and who met the study's screening procedures for schizophrenia (see "Measures of Independent Variables", DISSI). The *depression* group (n = 92) had an admitting diagnosis of major depression and met the study's screening procedure for that disorder. The *angina* group (n = 82) comprised patients admitted to university hospital medical units for evaluation or treatment related to their diagnosis of ischemic heart disease and who did not meet the study's screening criteria for schizophrenia or major depression.

Each hospitalized group comprised participants from two out of the three

sites participating in the study's data collection (Worcester, MA; Pittsburgh, PA; Kansas City, MO). Table 1 shows the proportional contributions of participants by each research site to each of the six groups and describes the groups demographically.

The psychiatric patients enrolled in this study represented 24% of patients with schizophrenia or schizoaffective disorder and 28% of patients with major depression (as defined by hospital clinicians' admitting diagnoses) who were admitted to the participating psychiatric units during the data collection period. The angina participants represented 12% of such admissions at the participating units.

"Attrition" prior to enrollment varied across sites. For example, requests by treating clinicians' that certain psychiatric patients not be approached for participation ranged from 1% to 20% of admitted patients in various hospitals. Additional patients were discharged before researchers had an opportunity to approach them. Because of these two sources of attrition, we were able to approach 43% to 56% of patients (depending on the site) who were eligible based on admitting diagnosis.

Among patients who were approached for participation, the rate for declining to participate varied across sites and diagnoses, averaging 15% to 20% of patients who were approached. Finally, an across-site average of 18% of psychiatric patients who agreed to participate eventually were excluded because their admitting diagnoses were not confirmed by the study's additional diagnostic screening procedure (DISSI: see "Measures of Independent Variables"). The study's screening criteria also excluded 32% of consenting patients with angina for whom one or the other of the two psychiatric diagnoses (usually depression) could not be ruled out.

Despite attrition, the demographic composition of the samples was similar to that of the admission populations from which they were drawn. Mean age, per-

Table 1. Site and Demographic Characteristics of Hospitalized and Community Groups

Variables	Schizophrenia		Depression		Angina	
	H[a]	C[a]	H	C	H	C
N						
Worcester	15	15	0	0	61	61
Kansas City	60	60	27	27	0	0
Pittsburgh	0	0	65	65	21	21
Total	75	75	92	92	82	82
Age						
M	34.8	34.1	35.0	34.7	55.3	55.3
SD	7.5	7.2	9.5	9.3	10.4	11.5
Percent male	52	52	45	45	64	64
Percent non-White	46	46	37	37	7	7
Socioeconomic status[b] (%)						
Levels I and II	1.3	0.0	4.3	3.3	19.5	24.4
Level III	18.7	22.7	31.5	33.7	29.3	32.9
Levels IV and V	80.0	77.3	64.2	63.0	51.3	42.7

[a] H = Hospitalized; C = community non-ill matched control.
[b] Socioeconomic status based on Hollingshead and Redlich (1957), using years of education and highest level of occupation attained since age 18. I = highest and V = lowest SES.

centage of racial minorities, and gender composition of the final samples were nearly identical to those of the pool of patients admitted to the hospital units with the relevant diagnoses during the study period. No method was available to determine differences in severity of disorder between the final samples and the pool of patients admitted to the hospitals during the study period. Clinicians' requests not to approach certain patients often were based on clinicians' judgments that the patients were too acutely disturbed to participate. Thus some of the most acutely mentally ill patients were not enrolled in the study.

Table 2 shows the status of the two mentally ill hospitalized groups on several patienthood variables. Current hospitalization was voluntary for most of the depression group and about one half of the schizophrenia group. The schizophrenia group had a somewhat younger mean age for first psychiatric admissions. Virtually all patients were undergoing treatment with psychoactive medications at the time they participated in the study. The majority participated in the study only a few days after admission (see Table 2); in many cases, therefore, research data were obtained from participants before one would expect substantial reduction of psychiatric symptoms as a result of treatment with medication.

Community Groups

The three community groups were drawn from the catchment areas of the psychiatric units participating in the study. They consisted of persons who did not meet the study's criteria for current schizophrenia, major depression, or ischemic

Table 2. Group Characteristics on Patienthood Variables

Variables	Schizophrenia (n = 75)	Depression (n = 92)	Angina (n = 82)
Legal admission status when tested (%)			
Emergency	17	14	0
Voluntary	57	80	100
Involuntary	28	5	0
Age first time ever hospitalized for this disorder (mean)	21.4	27.78	50.5
Number of Prior admissions (%)			
0–2	7	52	73
3 or more	84	45	27
No data	9	3	0
Current psychotropic medication (%)			
Antipsychotic	92	12	0
Antidepressants	6	34	0
Lithium	4	4	0
Benzodiazepines	57	17	0
Antiparkinsonian	61	17	0
Other	37	16	0
Days hospitalized at time of testing			
M	5.74	4.27	3.30
SD	7.4	2.5	2.2

heart disease (see "Procedure") and who reported that they never had been hospitalized for treatment of mental disorders or for cardiac conditions.

Each Community group represented a non-ill, never-hospitalized comparison group for one of the hospitalized groups, matched person-for-person on the following variables: age (within 5 years), gender, race, education (within 2 years), and highest lifetime occupation (within 1 increment on an 8-point scale of occupational level: Hollingshead & Redlich, 1957). Table 1 shows the results of the matching procedure, which was highly successful.

Independent Variables

Chart Data. Information obtained from hospital charts for the hospitalized participants included age, race, current (admitting) diagnosis, and date of admission. Additional information for hospitalized psychiatric patients included current medications, legal status (voluntary, involuntary, or emergency admission status at time of testing), and number of prior psychiatric hospital admissions.

Background Interview Data. A brief structured interview was used to obtain the following from all participants: number of prior psychiatric hospital admissions (if any), and current and past occupations (to obtain highest lifetime occupation). Additional information from psychiatric patients included age at first psychiatric hospitalization and number of years since first medication for the patient's current disorder.

Diagnostic Interview Schedule Screening Instrument (DISSI). The DISSI (Marcus, Robins, Bucholz, & Przybeck, 1989) uses items from the Diagnostic Interview Schedule to determine whether individuals meet criteria for selected DSM-III-R diagnoses. With psychiatric patients, the DISSI was used to confirm or disconfirm the presence of symptoms/signs that are required for the admitting diagnosis that qualified the patient to be approached for the study. It did not rule out the possibility that patients might meet criteria for other diagnoses as well. With angina and community participants, the DISSI was used to exclude participation by persons who met criteria for schizophrenia or major depression.

Brief Psychiatric Rating Scale (BPRS). The BPRS (Overall, 1988; Overall & Gorham, 1962; Overall & Klett, 1972) provides a method for rating the presence and severity of psychiatric symptoms/signs manifested by a psychiatric patient, based on a clinical interview during which inquiry and observation provide data for the ratings. The version of the BPRS used in this study contained 19 items.[1] Severity is rated on a 7-point scale, and overall severity of psychiatric condition is represented by the sum of the ratings. Past research has indicated that the BPRS is capable of yielding severity summary scores with substantial interrater reliability (Hedlund & Vieweg, 1980). BPRS summary scores above 40 commonly are associated with need for inpatient treatment.

[1] The BPRS used here was the 18-item version described in Overall (1988) and Overall and Klett (1972), plus one additional item ("elevated mood"). As explained by Overall (1988), this 19-item version was in use during the 1960s in the NIMH Psychopharmacology Service Center, although there is no publication reference for this version. It is very unlikely that research results would be different for one version than for the other.

Some analyses in this study employed scores on five BPRS factors identified in an earlier factor analysis of the BPRS (Guy, 1976; see also, Hedlund & Vieweg, 1980). Participants' ratings on items that loaded heavily on a factor were summed and divided by the number of items to produce factor scores. Each participant, therefore, obtained scores on each of five factors: I, Anxiety-Depression; II, Anergia; III, Thought Disorganization; IV, Activation; and V, Hostile-Suspiciousness.

Beck Depression Inventory (BDI). The BDI (Beck, 1978) is based on respondents' self-reports on 21 items referring to cognitive, affective, and vegetative symptoms/signs of depression. Scoring is objective; depression is considered to be *moderate* for scores of 19 to 29, and *severe* for scores of 30 or above.

Verbal Cognitive Functioning (VCF). Three subtests of the Wechsler Adult Intelligence Scale–Revised (WAIS-R) (Vocabulary, Similarities, Digit Span) were used to provide an index of participants' functioning on formal verbal cognitive tasks at the time of the research session. The VCF score was calculated by using WAIS-R norms to convert raw scores on subtests to scale scores, summing the three scale scores, multiplying by 2, and using WAIS-R age-normed tables to convert the calculation to the equivalent of a prorated Verbal IQ score. Though past research suggests that this method yields scores that are highly correlated ($r > .90$) with WAIS-R Verbal IQ (Wechsler, 1955), the present study employed the prorated score merely as an index of current verbal cognitive functioning. When used with psychiatric patients in acute phases of disorder, such measures cannot be presumed to be an index of "typical" or "baseline" intellectual functioning (or "intelligence"), because patients' formal cognitive functions may be impaired by current and transient symptoms of their acute psychiatric condition.

Measures of Dependent Variables

The study employed three instruments assessing four sets of abilities that are conceptually related to four legal standards for competence to make treatment decisions. The legal standards are described by Appelbaum and Grisso (1995) in a companion article to this one. A second companion article (Grisso, Appelbaum, Mulvey, & Fletcher, 1995) describes the abilities associated with these legal standards, as well as the development of the three instruments that were designed to assess those abilities. Collectively, they are called the "MacArthur Treatment Competence Research Instruments," copies of which are available upon request from the authors.

Understanding Treatment Disclosures (UTD: Grisso & Appelbaum, 1992) was designed to assess individuals' performance in a task that requires them to manifest *understanding* of information that is disclosed to them about a psychiatric disorder (schizophrenia or major depression) or a medical disorder (ischemic heart disease), as well as treatment options.[2] The UTD uses two modes of pre-

[2] The UTD is substantially the same instrument as an earlier one titled *Measuring Understanding of Disclosure* (MUD), which we developed for a preliminary study in preparation for the present one (Grisso & Appelbaum, 1991).

sentation of information: *uninterrupted disclosure*, involving presentation of five paragraphs (information about a disorder, a treatment, the benefits and risk/liabilities of the treatment, and an alternative treatment and its benefits and risks) followed by assessment of understanding of the information; and *element disclosure*, which involves presentation of the same information one paragraph at a time, followed by assessment of understanding after each element. The UTD procedure allows for assessment of understanding with two response modes: *paraphrase* recall, and *recognition*.

Together the stimulus and response modes form three sections (subtests) of the UTD: Uninterrupted–Paraphrase, Element–Paraphrase, and Element–Recognition. All questions assessing understanding are standardized, as is the above sequence of administration and the objective scoring of responses as presented in the UTD manual (Grisso & Appelbaum, 1992). Scores on each subtest range from 0 to 10. Evidence offered in a companion article (Grisso, Appelbaum, Mulvey, & Fletcher, 1995) indicates that the UTD can yield high interscorer reliability and has satisfactory internal consistency.

Perceptions of Disorder (POD: Appelbaum & Grisso, 1992) was designed to assess patients' beliefs related to the *appreciation* standard for competence to make treatment decisions. The POD's two subtests assess the degree to which patients fail to acknowledge the existence of disorders with which they have been diagnosed (Nonacknowledgment of Disorder: NOD), or fail to acknowledge the potential value of treatment for their disorders (Nonacknowledgment of Treatment Potential: NOT). Patients rate their degree of agreement or disagreement in response to statements about their disorder or potential benefit from treatment.

For the Nonacknowledgment of Disorder subtest, low scores (nonacknowledgment, failure of appreciation) are obtained only if patients fail to acknowledge symptoms that have been documented. For the Nonacknowledgment of Treatment Potential, low scores are obtained only if patients hold rigidly to beliefs that treatment will be of no benefit; assessment includes challenge with hypothetical questions designed to negate patients' original premises (see Appelbaum & Grisso, 1992; Grisso, Appelbaum, Mulvey, & Fletcher, 1995). The POD is a standardized procedure with objective scoring, yielding scores on each subtest ranging from 0 to 6. Lower scores indicate greater failure to acknowledge symptoms or potential treatment benefit.

Thinking Rationally About Treatment (TRAT: Grisso & Appelbaum, 1993) assesses the quality of individuals' *reasoning* about information in the process of deciding on treatment for a disorder. One part of the TRAT presents individuals with a third-person vignette involving information about several treatments for a disorder (schizophrenia, depression, or ischemic heart disease), including benefits and risks/discomforts of each treatment. Their explanations for their choice are scored according to the presence or absence of evidence for several problem-solving functions: information seeking, attention to the consequences of the treatments, making comparisons between various treatments, attention to the full range of treatments, and their ability to generate ideas about the potential everyday consequences of the treatments.'

Tasks in a second part of the TRAT assess individuals' abilities to reason

through problems involving transitive proposals and probability statements and the consistency with which they use personal preferences in choosing among alternatives. (See Grisso & Appelbaum, 1992, and Grisso et al., 1995, for more detailed descriptions of the TRAT subtests and reliability of measurement.)

Factor analyses of the TRAT subtests (Grisso, Appelbaum, Mulvey, & Fletcher, 1995) indicated greater internal consistency for a TRAT measure based on six of the eight subtests. Therefore, several analyses in the present study were performed for both the eight-subtest version, called *TRAT-1*, and the six-subtest version, called *TRAT-2*. Scores for all subtests are obtained according to objective criteria described in the TRAT manual (Grisso & Appelbaum, 1993). The TRAT-1 yields scores of 0 to 19, and the TRAT-2 range is 0 to 14.

Expressing a Choice (EC) measured with a single item in the TRAT, assesses whether individuals are capable of selecting and stating a choice, or whether they manifest ambivalence or confusion that results in a failure to reach a conclusion about a preferred treatment. EC scores range from 0 to 2.

Reliability of Measurement

Data were collected by 10 research assistants in three cities. Several procedures were used to maximize consistency of data collection across sites. All research assistants received initial training in all measures at a 2-day joint training conference held at the central project site (Worcester, MA). The principal investigators and project director then engaged in a series of training sessions at the various sites, including sessions in scoring and observation of research assistants in practice administrations with patients. BPRS and DISSI administration and scoring were practiced under supervision in small groups until consensus appeared to be routine.

Interscorer reliability for the dependent measures is reported in a companion article (Grisso et al., 1995) and was based on protocols that were scored independently by various research assistants, then compared within sites and to a master scorer at the central project site. Spot checks for scoring were performed periodically during the process of data collection, which required about 15 months.

Procedure

Hospitalized Samples

Admissions logs on participating hospital units at the three sites were examined on each working day for patients admitted with diagnoses and ages (18–70) appropriate for the study's design. Age, gender, and race of those patients were recorded, and we solicited their treating clinicians' permission to approach the patients for possible research participation. Reasons for clinicians' denial of permission were recorded. We often reapproached clinicians a few days later to determine whether a patient's status had changed sufficiently to allow participation. According to the study's design, however, patients were ineligible for enrollment after 14 consecutive days of hospital stay.

When clinician clearance had been obtained, researchers explained to those

patients the nature of the study, the tasks they would be asked to perform, confidentiality of participants' responses, and the distinction between the researchers and hospital personnel involved in patients' treatment. They were offered $10 for their participation. (Patients who declined to participate generally were reapproached one more time within 2 or 3 days.) Patients who expressed interest were provided a more complete description and written consent was obtained.

For patients who consented, research sessions were conducted in interview rooms on hospital wards or at bedside for some ischemic heart disease patients. Researchers first administered the DISSI, obtained background interview data, and obtained relevant data from participants' hospital charts. Psychiatric patients whose DISSI results did not confirm their chart diagnosis of schizophrenia or major depression and angina patients whose DISSI did not rule out these two diagnoses were given the $10 for their participation and did not participate in the remainder of the study's procedures. Patients who met the study's DISSI criteria were then administered the BPRS (schizophrenia and depression samples only), BDI, WAIS-R subtests contributing to the VCF, and the UTD, POD, and TRAT. Participants received the forms of the latter three instruments that matched their own current diagnosis.[3]

Community Samples

Participants for the community samples were recruited in several ways, both within and across data collection sites. The majority were obtained in response to advertisements placed in local newspapers. In addition, notices soliciting participants were posted in a variety of facilities and shops (e.g., homeless shelters, barber and beauty shops, churches, fast-food, and other commercial establishments) located in neighborhoods where patients in the hospitalized psychiatric samples were known to reside. This was intended to attract research participants with socioeconomic characteristics similar to the hospitalized samples.

Individuals who responded to advertisements by telephoning the research office were asked to provide their age, gender and race, occupation, and years of education to determine eligibility for the study. Early in the study, appointments were made to meet these respondents whenever they were within the study's age range (18–70) and spoke English sufficiently well to participate. As the pool of hospitalized participants grew, acceptance of respondents from the community became more selective, based on the need to find community participants who could be matched with hospitalized participants on age, gender, race, and socioeconomic status.

Data were collected in an interview and testing session held at participants' homes or in a public place that would assure an adequate degree of researcher safety and a minimum of distractions (e.g., a room in a public library, a booth in a fast-food restaurant at a low-volume time of day). Sessions began with a process of informed consent similar to the procedure with hospitalized patients. Partici-

[3] A preliminary study (Grisso & Appelbaum, 1991) indicated that individuals' performances are relatively consistent across various forms of the instruments.

pants were screened with the depression and schizophrenia modules of the DISSI, and those meeting either set of criteria were excluded from the study. Background interview data, BPRS, BDI, and WAIS-R (VCF) data were collected, followed by administration of the UTD and the TRAT. (The POD was not administered to community participants, because it is not applicable to persons who have no disorder.) Each community participant received the same diagnostic version of the UTD and TRAT as the hospitalized participant to whom he or she was matched demographically and socioeconomically. Community participants also were compensated financially for the session.

Test–Retest Procedure

Subsamples of the schizophrenia, depression, and community samples were readministered the BPRS, BDI, WAIS-R subtests, and dependent measures between 14 and 20 days after first administration. Details of this procedure and its results are reported elsewhere (Grisso et al., 1995).

Data Processing

Research assistants scored and coded all data obtained from participants they had tested. A project director at each site was responsible for monitoring accuracy of scoring and coding. The project director at the main research site reviewed coded forms from all sites and rescored protocols for the dependent measures on a random basis. A "master scorer" at the main research site reviewed and rescored all TRAT protocols from all sites, because examination of interscorer reliability with the TRAT (Grisso, Appelbaum, Mulvey, & Fletcher, 1995) indicated less than high reliability in scoring by some research assistants for some TRAT subtests.

RESULTS

Mental Status

Table 3 describes the hospitalized and community groups' scores on the BPRS (severity of psychiatric symptoms), BDI (severity of depression), and VCF (WAIS-R index of verbal cognitive functioning). Both of the hospitalized groups with mental illness scored significantly lower on the VCF index than their matched community groups. The schizophrenia group scored significantly higher than its Community group on the BPRS and BDI, and the depression group scored significantly higher than its matched community group on the BDI, but not on the BPRS. The angina group had a slightly (but significantly) higher BDI score than its community group.

The BDI mean scores and standard deviations for the Depression group and its community comparison group indicate that the depression participants generally scored in the "high moderate" and "low severe" range of depression, whereas the community participants scored well below the "mild" range. The

Table 3. Mental Status Measures

Variables	Schizophrenia (n = 75)		Depression (n = 92)		Angina (n = 82)	
	H[a]	C[a]	H	C	H	C
Brief Psychiatric Rating Scale (BPRS)						
M	46.28	27.93	32.35	28.48	NA	NA
SD	10.6	6.7	8.6	6.8		
Beck Depression Inventory (BDI)						
M	21.75	6.60	29.99	5.50	7.95	6.04
SD	14.8	5.5	11.4	5.3	5.7	5.5
Verbal Cognitive Functioning (VCF: Verbal WAIS-R)						
M	79.93	90.25	89.28	96.56	97.58	103.20
SD	11.8	13.5	13.7	11.4	11.6	12.2

[a] H = Hospitalized; C = community non-ill matched control. Schizophrenia H vs. C means were significantly different at the .0001 level of significance for all three mental status measures. Depression H vs. C means were significant for the BDI ($p < .0001$) and the VCF measure ($p < .05$). Angina H vs. C means differed significantly only on the BDI ($p < .05$).

lack of difference between their mean scores on the BPRS, therefore, was not interpretable as a lack of difference between the groups in severity of depression. Instead, it appeared to be due to consistently low ratings on the BPRS by researchers in one of the two sites that provided data for the depression group. That site produced a mean score for depression participants on the BDI (28.72) that was very similar to that of the second site (32.92), while its mean BPRS score (for which the response format is not as clearly anchored as the BDI) was much lower (27.94) than the second site's (42.97). Because of these results, the measure with the more structured response format, the BDI, was used for all further analyses requiring an index of symptom severity for the depression group.

As seen in Table 3, the VCF scores (prorated verbal score based on three WAIS-R subtests) of the schizophrenia group were significantly lower than those of their matched community group. We interpret this as reflecting the effects of acute thought disorder on some of the schizophrenia patients' performance, which should not be interpreted as a difference in premorbid or "true" intellectual capacity.

Understanding Treatment Disclosures (UTD)

Group Performance

Means and standard deviations for the 6 groups on the UTD (understanding) are shown in Table 4. The schizophrenia group scored significantly lower than its community group on all 3 UTD subtests. The depression group scored significantly lower than its community group only on the Element–Paraphrase and Element–Recognition subtests. There were no significant differences between the angina group and its matched community group on any of the UTD subtests.

The 3 hospitalized groups were compared to each other on each UTD subtest with analyses of covariance that controlled for group differences in age, highest

Table 4. Means on Subtests of the UTD ("Understanding")

UTD subtests	Schizophrenia (n = 75)		Depression (n = 92)		Angina (n = 82)	
	H[a]	C[a]	H	C	H	C
Uninterrupted—Paraphrase						
Mean	4.74	6.09	6.57	6.69	6.65	6.61
SD	3.0	2.9	2.7	2.7	2.3	2.7
t, *p*	3.01,	$p < .003$	n.s.		n.s.	
Element—Paraphrase						
M	6.18	8.40	8.01	8.72	8.35	8.08
SD	3.0	1.7	2.0	1.4	1.9	2.0
t, *p*	5.50,	$p < .0001$	2.83,	$p < .005$	n.s.	
Element—Recognition						
M	6.89	8.38	8.72	9.22	8.75	8.96
SD	2.8	1.7	1.7	1.1	1.2	1.6
t, *p*	3.82,	$p < .0001$	2.31,	$p < .02$	n.s.	

[a] H = Hospitalized; c = Community non-ill matched control.

occupation, and education.[4] Results indicated that the groups were significantly different in their performance on all three subtests: Uninterrupted–Paraphrase, $F(2,246) = 8.03, p < .0001$; Element–Paraphrase, $F(2,246) = 10.48, p < .0001$; Element–Recognition, $F(2,246) = 15.74, p < .0001$. Additional *t*-tests indicated that the differences on each of the subtests derived from statistically significant differences between the schizophrenia and depression groups (two-tailed, $p < .005$ on all subtests) and between the schizophrenia and angina groups (two-tailed, $p < .002$ on all subtests). There were no significant differences between the depression and angina groups on any of the UTD subtests.

The distributions of Uninterrupted–Paraphrase scores for the hospitalized and community groups are shown in Figures 1 and 2; Figures 3 and 4 provide similar comparisons for Element–Paraphrase scores. Whereas scores were skewed to the left for most groups, scores for the schizophrenia group were skewed to the right on Uninterrupted–Paraphrase and approximated a normal distribution on Element–Paraphrase. On Element–Recognition (no figure provided), the schizophrenia group manifested a distribution skewed to the left. However, 31% of the schizophrenia group continued to obtain scores of 5 or lower on this subtest, compared to only 5% of the depression group and 2% of the angina group.

[4] The comparisons between hospitalized groups were considered secondary analyses, the primary comparisons being between hospitalized groups and their matched non-ill community controls. Differences between the hospitalized groups in age, race, and socioeconomic status were substantial (see Table 2). Preliminary analyses indicated that scores on several of the treatment competence instruments were related significantly to age and socioeconomic status. Therefore, we used age, education, and highest occupation as covariates in all secondary comparisons of hospitalized groups.

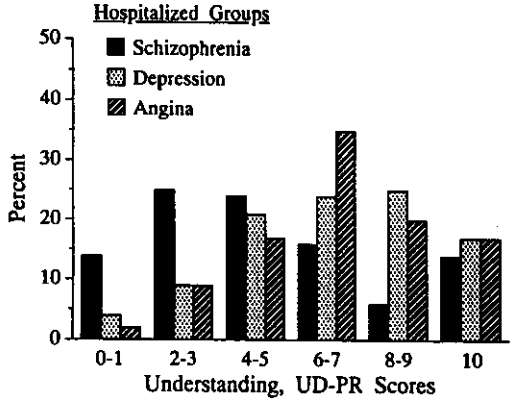

Fig. 1. Hospitalized groups, understanding on UTD uninterrupted disclosure—paraphrase response (UD-PR).

Relation of UTD Performance to Other Variables

As shown in Table 5, UTD subtest scores among the hospitalized participants with mental illness were correlated significantly (negative relation) with severity of symptoms for the schizophrenia group only, while verbal cognitive functioning was correlated significantly (positive relation) with UTD performance for the depression group. For both mentally ill groups, socioeconomic status was correlated significantly (positive relation) with performance on the subtests that required paraphrased responses.

To examine further the significant relation between BPRS scores and UTD performance in the schizophrenia group, we calculated correlations between UTD subtests and five BPRS factor scores (Guy, 1976). Using Pearson r, scores on BPRS Factor III (Thought Disorganization) were correlated significantly with scores on each of the UTD subtests (Uninterrupted–Paraphrase, $r = -.35$, $p < .001$; Element–Paraphrase, $r = -.44$, $p < .0001$; Element–Recognition, $r =$

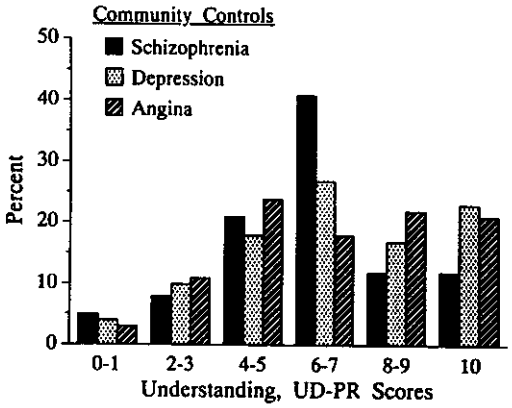

Fig. 2. Non-ill community control groups, understanding on UTD uninterrupted disclosure—paraphrase response (UD-PR).

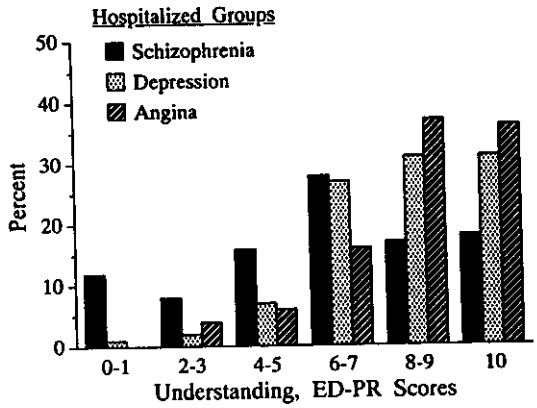

Fig. 3. Hospitalized groups, understanding on UTD element disclosure—paraphrase response (ED-PR).

$-.39$, $p < .0001$). No other significant correlations were found between BPRS factors and UTD subtest scores.

Table 6 shows the relation between BPRS Factor III scores and scores on the Element–Paraphrase subtest of the UTD, for patients in the schizophrenia group scoring in the low (0–4), median (5–8), and high (9–10) ranges of this subtest. These data suggest that the relation of poorer understanding on the UTD to greater Thought Disorganization (Factor III scores) was manifested especially in three BPRS items: Conceptual Disorganization, Unusual Thought Content, and to some extent, Hallucinatory Behavior.

Perception of Disorder (POD)

Group Performance

Table 7 shows mean scores for the three hospitalized groups on the two subtests of the POD ("appreciation"). About one-third of schizophrenia patients

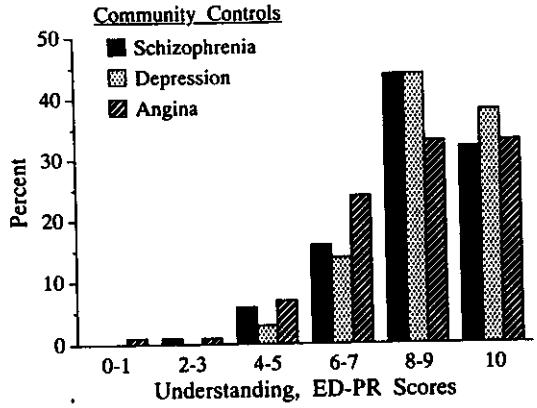

Fig. 4. Non-ill community control groups, understanding on UTD element disclosure—paraphrase response (ED-PR).

Table 5. Correlations Between UTD ("Understanding") Scores and Mental Status and Patienthood Variables

Variables	Schizophrenia (n = 75)			Depression (n = 92)		
	UD-PR	ED-PR	ED-RC	UD-PR	ED-PR	ED-RC
Symptom severity[a]	−.33*	−.41*	−.41*	−.03	.01	−.06
Verbal cognitive functioning	.12	.26	.33*	.46*	.39*	.34*
Prior admissions	−.03	.01	.02	−.05	−.02	.00
Age first hospitalization	.08	.08	.17	.11	.11	.00
SES	.27*	.35*	.24	.31*	.27*	.35*

* $p < .01$.
[a] BPRS for schizophrenia group, BDI for depression group. UD-PR, uninterrupted disclosure—paraphrase response; ED-PR, element disclosure—paraphrase response; ED-RC, element disclosure—recognition.

obtained Nonacknowledgment of Disorder subtest scores of 3 or lower (tended not to acknowledge some aspect of their disorders), while this was found for only a very small minority within the other two groups. An analysis of covariance (controlling for age, highest occupation, and education) revealed significant differences between groups on this subtest: $F(2,246) = 14.30, p < .0001$. Additional t-tests comparing pairs of groups indicated that the performances of the patients in the depression and angina groups were not significantly different from each other, but both were significantly more likely than the patients in the schizophrenia group to acknowledge their disorder: schizophrenia vs. angina, two-tailed, $p < .0001$; schizophrenia vs. depression, two-tailed, $p < .0001$.

A different pattern was found on the Nonacknowledgment of Treatment Potential subtest (see Table 7). About 13% of schizophrenia patients and 14% of

Table 6. BPRS Factor III Ratings for Subgroups of Schizophrenia Sample Scoring in Low, Median, and High Ranges of Element–Paraphrase Subtest ("Understanding")

BPRS variables	Scoring subgroups on element–paraphrase		
	Low (0–4) (n = 20)	Median (5–8) (n = 35)	High (9–10) (n = 20)
Factor III Ratings			
M	14.60	11.21	9.85
SD	10.7	10.2	8.6
Percent with ratings equal to or above 4 on all Factor III symptoms	45.0	23.5	5.0
Percent with ratings equal to or above 4 on Factor III items			
Conceptual Disorganization	65.0	32.4	15.0
Unusual Thoughts	65.0	52.9	35.0
Grandiosity	45.0	29.4	20.0
Hallucinations	70.0	38.2	55.0

Table 7. POD Scores ("Appreciation") for Hospitalized Groups

POD subtests	Schizophrenia (n = 75)	Depression (n = 92)	Angina (n = 82)
Nonacknowledgment of Disorder			
M	3.96	5.54	5.37
Percent with scores of 3 or below	34.7	4.4	9.8
Nonacknowledgment of Treatment Potential			
M	4.89	5.00	5.80
Percent with scores of 3 or below	13.0	14.1	0.0

Note: Scores on both of the POD subtests range from 0 to 6, with lower scores representing failure to acknowledge one's disorder or the potential value of treatment.

patients in the depression group obtained scores of 3 or lower on this subtest (tended not to acknowledge the potential value of treatment). Analysis of covariance indicated that these results were statistically significant: $F(1,246) = 3.07$, $p < .04$. Additional t-tests between pairs of groups indicated that there was no significant difference between the schizophrenia and depression groups, but both of these groups were significantly less likely than the angina patients to acknowledge the potential value of treatment (schizophrenia vs. angina, two-tailed, $p < .005$; depression vs. angina, two-tailed, $p < .005$).

Relation of POD Performance to Other Variables

For both the schizophrenia and depression groups, there were no significant correlations between scores on either of the POD subtests and scores on any other mental status or patienthood variable (BPRS total and factor scores, BDI, verbal cognitive functioning, SES, number of prior hospital admissions, and age at first admission). It was noted on Nonacknowledgment of Disorder, however, that the percentage of schizophrenia patients who obtained high ratings (4 or above) on BPRS "Conceptual Disorganization" was greater for those who tended not to acknowledge their disorder (50%) than for those who fully acknowledged their disorder (12%). On the Nonacknowledgment of Treatment Potential, the percentage of schizophrenia patients with chart diagnoses of paranoid schizophrenia was greater for those who failed to acknowledge the potential value of treatment (75%) than for those with partial (55%) or full (48%) acknowledgment. The most frequent reasons given by schizophrenia patients for devaluing treatment, especially medication, involved beliefs that it was intended to harm them in some way. In contrast, the most frequent reason given by depression patients for devaluing treatment was the belief that they were "too sick" for anything to help them.

Thinking Rationally About Treatment (TRAT)

Group Performance

Table 8 shows means and standard deviations for the 6 groups for total scores on the TRAT ("reasoning"), separately for TRAT-1 and TRAT-2. (As noted ear-

Table 8. Means on TRAT-1 and TRAT-2 ("Reasoning")

TRAT versions	Schizophrenia (n = 75)		Depression (n = 92)		Angina (n = 82)	
	H[a]	C[a]	H	C	H	C
TRAT-1 (range 0–19)						
M	10.12	13.68	12.83	14.25	14.15	14.31
SD	4.1	3.4	3.2	2.6	2.6	2.5
t, p	5.70,	$p < .0001$	3.20,	$p < .001$	n.s.	
TRAT-2 (range 0–14)						
M	7.86	10.57	10.01	10.97	10.81	10.86
SD	3.2	2.8	2.7	2.1	2.0	2.2
t, *p*	5.50,	$p < .0001$	2.72,	$p < .007$	n.s.	

[a] H = Hospitalized; C = community non-ill matched control.

lier, TRAT-2 is based on 6 of the 8 TRAT subtests, providing a TRAT with greater internal consistency.) Distributions of TRAT-2 scores (which were quite similar to those for TRAT-1 scores) are shown in Figures 5 and 6. Both the schizophrenia and depression groups scored significantly lower than their matched community groups. There was no significant difference between the angina group and its matched community group on TRAT-1 or TRAT-2.

Scores of the 3 hospitalized groups were compared on TRAT-2 with analysis of covariance that controlled for group differences in age, highest occupation, and education. Results indicate that the groups were significantly different in their performance: F (2,246) = 14.55, $p < .0001$. Additional t-tests indicate that this result reflected significant differences between the schizophrenia and depression groups (two-tailed, $p < .001$), the schizophrenia and angina groups (two-tailed, $p < .0001$), and between the depression and angina groups (two-tailed, $p < .02$).

Table 9 shows the means for each group on each of the TRAT subtests. The extent of group differences is seen more clearly, however, by examining the percentage of participants in each group who received low credit on a subtest (0 credit on subtests with a 0–2 range, and 0–1 credit for subtests that had a 0–3 range).

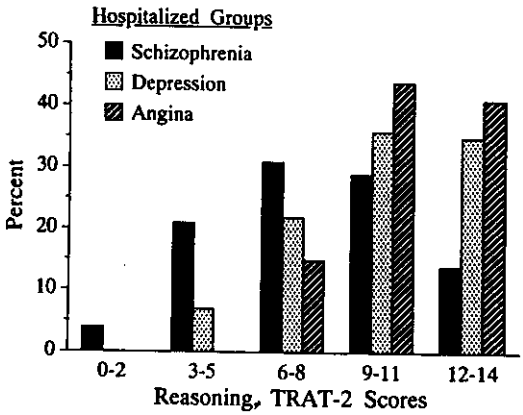

Fig. 5. Hospitalized groups, reasoning on TRAT-2.

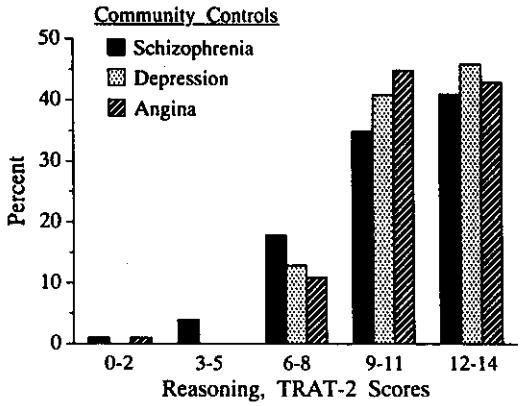

Fig. 6. Non-ill community control groups, reasoning on TRAT-2.

An additional series of *t*-tests indicated that the schizophrenia group scored significantly more poorly than its matched community group on six of the eight subtests, while depression patients were significantly lower than their community matches on four subtests.

Table 9. Means on TRAT ("Reasoning") Subtests, and Percent of Participants Scoring in Lowest Category[a]

	Schizophrenia (n = 75)			Depression (n = 92)			Angina (n = 82)	
TRAT subtests[b]	H		C	H		C	H	C
Means								
CONS	1.16	*	1.46	1.34		1.40	1.42	1.36
COMP	0.74		1.05	1.03		0.97	1.02	1.89
COMX	1.00	***	1.50	1.63		1.70	1.59	1.48
GENER	1.28		1.54	1.37	*	1.62	1.64	1.78
TRANS	1.91	***	2.46	2.24	***	2.63	2.62	2.72
PROBA	1.77	***	2.53	2.39	**	2.64	2.50	2.63
SEEK	0.76	**	1.22	0.85	*	1.16	1.22	1.46
WEIGH	1.51	*	1.88	1.95		2.15	2.09	1.98
Percent in lowest category								
CONS	13.3	*	2.7	7.6		3.3	11.0	9.8
COMP	49.3		33.3	30.4		29.3	32.9	42.7
COMX	38.7	***	12.0	10.9		6.5	12.2	13.4
GENER	24.0		13.3	18.5		13.0	11.0	7.3
TRANS	32.4		14.7	16.3		3.3	3.7	4.9
PROBA	42.7	**	10.7	16.3		2.2	8.5	3.7
SEEK	61.3	**	36.0	56.5	*	39.1	39.0	26.8
WEIGH	48.6	**	36.0	25.0		14.3	18.5	29.6

[a] Lowest category is score of 0 on subtests with 0–2 range, and scores of 0–1 on subtests with 0–3 range.

[b] CONS = Consequential thinking; COMP = comparative thinking; COMX = complex thinking; GENER = generating consequences; TRANS = transitive thinking; PROBA = probabilistic thinking; SEEK = seeking information; WEIGH = weighting consequences.

* $p < .01$.

** $p < .005$.

*** $p < .0001$.

Relation of TRAT Performance to Other Variables

As shown in Table 10, TRAT-1 and TRAT-2 scores for the hospitalized mentally ill participants were correlated significantly only with verbal cognitive functioning (positive relation). Only TRAT-1 scores for the schizophrenia group were correlated significantly (negative relation) with SES. Despite the lack of a significant relation between BPRS scores and TRAT performance in the schizophrenia group, correlations were examined between TRAT-2 and the five BPRS factor scores noted earlier (Guy, 1976). Trends were apparent similar to those seen in the relation described earlier between UTD subtests and BPRS factor scores (i.e., for "Thought Disorganization"), but the relations were not statistically significant.

Performance Across Measures

In a companion article (Grisso et al., 1995) we reported that scores on UTD subtests were modestly correlated with TRAT scores, but scores on both of these measures correlated poorly with POD subtests scores. These earlier correlational analyses suggested the need to examine more closely the cross-measures performance of patients. To do so, we classified participants according to their impaired and adequate performance on each measure, and examined the frequency of intersections of impaired and adequate performance across the three dependent measures for each of the 6 groups. *Impaired understanding* was defined as a UTD Element–Paraphrase score of 0–4, and *impaired reasoning* was defined as a TRAT-2 score of 0–5. These criteria were set at a level that would include less than 5% of any of the community samples. *Impaired appreciation* was defined as performance resulting in the lowest possible score (0) on the Nonacknowledgment of Disorder and/or 0–2 on Nonacknowledgment of Treatment Potential, an extremely conservative cutoff.

Table 11 shows the percentages of each sample classified as having manifested impaired performance on each of the measures. The schizophrenia group had much higher proportions than the other two hospitalized samples in the impaired categories on all measures, comprising about one-quarter of the schizophrenia sample on any given measure. To interpret these results further, we ex-

Table 10. Correlations Between TRAT ("Reasoning") Scores and Mental Status and Patienthood Variables

	Schizophrenia (n = 75)		Depression (n = 92)	
Variables	TRAT-1	TRAT-2	TRAT-1	TRAT-2
Symptom severity[a]	−.25	−.25	−.10	−.08
Verbal cognitive functioning	.39*	.37*	.27*	.30*
Prior admissions	−.16	−.18	−.14	−.06
Age first hospitalization	.25	.07	−.03	−.07
SES	.28*	.23	.18	.22

[a] BPRS for schizophrenia group, BDI for depression group.
* $p < .01$.

Table 11. Percent of Cases Scoring in "Impaired"[a] Range on Dependent Measures

Dependent measures	Schizophrenia (n = 75)	Depression (n = 92)	Angina (n = 82)	Community (n = 249)
UTD/Element—Paraphrase (understanding)	28.0	5.4	7.3	2.4
TRAT-2 (reasoning)	24.0	7.6	0.0	2.0
POD (appreciation)	22.6	11.9	2.4	NA

[a] "Impaired" criteria were: understanding, scores of 0–4 on UTD/Element—Paraphrase; reasoning, scores of 0–5 on TRAT-2; appreciation, score of 0 on Nonacknowledgment of Disorder and/or 0–2 on Nonacknowledgment of Treatment Potential.

amined participants' patterns of "impairment" across all three measures together. Consistent with the earlier finding of modest to low correlations between the three instruments (Grisso et al., 1995), a substantial number of participants showed impaired performance on some measures while performing adequately on others. For example, while about three-quarters of the schizophrenia participants performed adequately on any one measure, only 48.1% of them manifested adequate performance across all three measures. This compared to 76.1% of the depression group, 87.8% of the angina group, and 96.0% of the three community control groups combined.

Expressing a Choice

The ability to express a choice among treatment options was assessed by an item in the TRAT vignette, asking respondents which of the three treatments they would recommend to the hypothetical person in the vignette. Only 5% of the respondents were unable to state a preference (i.e., scored 0 on the item), usually because of extreme ambivalence. This included 2.7% of the schizophrenia group, 7.6% of the depression group, and 3.7% of the angina group. There were no statistically significant differences between the hospitalized groups and their community control groups on this item, although the schizophrenia group's Community controls included more cases (12.2%) of failure to state a preference than the schizophrenia group.

DISCUSSION

The purpose of this study was to offer a reliable, empirical perspective for legal, policy, and clinical decisions regarding the potential for incompetence to make treatment decisions among patients with mental and medical illnesses. The dependability of the results is supported especially by (a) the demonstrated reliability of the measures of legally relevant decisional abilities (Grisso et al., 1995), (b) the multisite sampling strategy and careful efforts to assure uniformity in procedure across sites, and (c) the study's comparisons of patients to non-ill control groups matched with them on critical demographic variables. The relevance of the data for legal and policy interpretations is enhanced especially by (a)

the development of multiple experimental measures of ability in accord with conceptual analyses of the legal standards associated with competence to consent to treatment, and (b) the use of comparison groups that represent the psychiatric populations of greatest concern for policy and clinical debates.

Summary of Main Findings

There were three main findings in this study. *First,* on the measures of understanding (UTD), appreciation (POD), and reasoning (TRAT), as a group, patients with mental illness more often manifested deficits in performance than did medically ill patients and their non-ill control groups. Indeed, when the most highly impaired subgroups were identified on each measure, they were composed almost entirely of patients with mental illness.

Second, despite overall lower levels of performance in the groups with mental illness, there was considerable heterogeneity within and across the schizophrenia and depression groups. Impairments in performance were more pronounced and more consistent across measures for the schizophrenia patients than for patients with depression. This finding is consistent with a large body of research establishing the poorer performance of patients with schizophrenia, compared to normal controls, on a wide array of cognitive tasks (Gold & Harvey, 1993). Even so, on any given measure of decisional abilities, the majority of patients with schizophrenia did not perform more poorly than other patients and nonpatients. The poorer mean performance of the schizophrenia group for any particular measure was due to a minority within that group.

Third, among patients with schizophrenia, the minority with poorer performance on the measures of understanding and reasoning tended to manifest greater severity of psychiatric symptoms, especially those of thought disturbance (e.g., conceptual disorganization, unusual thoughts). These results are in keeping with both theory and empirical findings regarding cognitive deficits associated with schizophrenia (e.g., Clare, McKenna, Mortimer, & Baddeley, 1993). Apart from this difference, however, this poorer-functioning subgroup was not distinguishable on the basis of other demographic, mental status, or patienthood variables used in this study.

Limits of Interpretations

Interpretations of these findings for legal, policy, and clinical decisions should acknowledge two important conceptual matters in order to avoid misinterpretation or misuse of the findings.

Underestimation of Deficits Due to Sampling for Patient Groups

The findings probably underestimate the proportion of patients (with both mental and medical illnesses) who have serious deficiencies in decisional abilities related to competence to consent to treatment. As described in the Method section, some patients were not enrolled as research participants because their doctors believed that they were too acutely disturbed to participate. Many of those

patients, had they been included in the study, probably would have manifested very poor performance on the measures of decisional abilities. The proportion "screened out" by treating professionals was not systematically different for the various diagnostic groups; the proportion seemed to vary more by hospital site than by diagnosis. Therefore, the degree of the underestimation cannot be specified in our data.

Conceptual Differences Between Ability Measures and Determinations of Legal Competence

The experimental measures developed for this study should not be interpreted as though they provide determinations of legal incompetence to consent to treatment. This would be inappropriate for several reasons.

First, although empirical data do not exist to demonstrate it, judges undoubtedly differ in their applications of legal standards when making competence determinations. In addition, as we described earlier (Appelbaum & Grisso, 1995), statutes or legal precedent may require judges in different jurisdictions to attend to different legal standards (that is, only one of the four legal standards, or some combination of them) and therefore different types of decisional abilities among those that were measured in this study. Therefore, any attempt to fashion a proxy determination of legal incompetence based on the results of this study would lack meaning across jurisdictions.

Second, determinations of incompetence require a judgment that the *degree* of deficits in the abilities relevant in a particular case is sufficiently great to warrant a declaration of incompetence, with consequent invalidation of the person's choice. There is no numerical criterion that can represent this judgment across cases, because the degree of deficits in ability that logically will be required may be expected to vary in relation to the specific disorder, proposed treatments, probable consequences, and other contextual factors that vary from one case to another (see the "interactive" quality of legal competencies: Grisso, 1986). In contrast, the criterion scores used in this study to define "impaired" and "adequate" performance were based on the relative infrequency of scores below a particular point, not on presumptions regarding the adequacy or inadequacy of performance as it relates to requirements for making treatment decisions.

On the other hand, we believe that the measures do provide meaningful representations of the decisional abilities that courts have considered when making competence determinations. Moreover, it seems likely that most persons who are determined to be legally incompetent to consent to treatment would perform poorly on these measures. Yet many low scorers might not be found incompetent. In other words, we suspect that the measures may have reasonable sensitivity in that few persons that any court would find to be legally incompetent would perform well on all of the measures; but specificity is uncertain in that low scores on one or more of the measures might not be highly predictive of legal determinations of incompetence. Low scorers, therefore, might best be seen as "at greater risk" of failing to meet thresholds of ability associated with determinations of legal competence.

Policy Implications for Patients' Rights to Give or Withhold Consent for Treatment

Whether persons with mental illness should be permitted to give or withhold consent for treatment to the same extent and in the same manner as persons with medical disorders has been the subject of considerable controversy in the courts (e.g., *Rogers v. Commissioner,* 1983; *Zinermon v. Burch,* 1990). Many of the arguments offered in these debates turn on premises concerning the degree to which persons with mental illness manifest impairment of their decision-making abilities (Appelbaum & Grisso, 1995). The data from this study should assist in assessing the validity of these arguments.

Opponents of allowing persons with mental illness equal decision-making rights no longer can maintain that *all* persons who are in need of hospitalization for mental disorder lack the requisite abilities to make decisions regarding their treatment. Nearly one half of the schizophrenia group and 76% of the depression group performed in the "adequate" range (according to *ad hoc* definitions of adequacy used in this study) across all decision-making measures, and a significant portion performed at or above the mean for persons without mental illness. When performance on a single measure is examined, as may be relevant, for example, in jurisdictions that have adopted only an understanding standard for legal competence, the rate of adequate performance rose to roughly 75% for patients with schizophrenia and to approximately 90% or more for patients with depression. Thus, the justification for a blanket denial of the right to consent to or refuse treatment for persons hospitalized because of mental illness cannot be based on the assumption that they uniformly lack decision-making capacity.

On the other hand, the data confirm that significant differences exist in decision-making abilities between persons with and without mental illness, especially when the comparisons focus on patients with schizophrenia. For any given measure, approximately 25% of the schizophrenic group scored in the "impaired" range, compared with 5%–7% of angina patients and only 2% of community controls. Some patients were in the "impaired" range on one measure but not on another. When all measures are combined, 52% of patients with schizophrenia showed impairment on at least one measure, in contrast to 12% of angina patients and 4% of community controls. The high frequency of deficiencies in decision-making abilities in patients with schizophrenia, and to a lesser extent in patients with depression, may justify differences in the ways in which consent to treatment for mental illness and for medical illness are handled.

In medical settings, patients generally are presumed competent to consent to treatment, and their choices are honored. Questions concerning patient competence tend to be raised only when patients decline to follow physicians' recommendations, and even then only in a small minority of cases (Appelbaum & Roth, 1982). Clinicians in psychiatric settings behave similarly (Lidz et al., 1984), usually challenging patients' competence only in those cases in which prescribed treatment is declined (and not routinely even then: see Hoge et al., 1990). The law, too, has focused heavily on treatment refusal in creating special rules for dealing

with patients whose decision-making abilities may be impaired (Appelbaum, 1988).

Assuming a goal of policy in this area is to protect the interests of persons with mental illness who are unable to act on their own behalf, a primary focus on patients who refuse treatment omits consideration of the needs of the majority of impaired patients. The largest study of treatment refusal in psychiatric facilities revealed that approximately 7% of patients hospitalized for mental illness refuse treatment (Hoge et al., 1990), a figure comparable to those in most other reports. In contrast, our data indicate that at least 23%–52% of patients hospitalized with schizophrenia and 5%–24% of patients hospitalized with major depression may have substantially impaired decision making. Policies targeted only at the small percentage of patients who refuse treatment fail to meet the needs of the larger group of patients who, although impaired in decisional abilities, assent to treatment.

Whether special efforts should be made to detect patients with severely reduced decision-making abilities depends on the costs of such a program and the benefits likely to be achieved (Appelbaum & Schwartz, 1992). These considerations go well beyond the scope of this article. The high base rates of impairment that our data suggest, however, indicate that if general screening for decision-making deficiencies were believed to be important, the process could be conducted efficiently. This would be so particularly if groups at special risk were targeted, such as those patients with schizophrenia who manifest significant thought disorder. Once detected, these patients need not be deprived automatically of decision-making rights. They might receive extra educational efforts to improve their performance, or additional protection, including independent clinical review, could be used to ensure appropriateness of care. The latter may be particularly useful given data suggesting that patients with impaired decision-making abilities run a heightened risk of receiving suboptimal care (Hoge, personal communication).

Caution is required, though, in proposing routine resort to the courts in dealing with this problem. Were every patient with impaired capacities to require judicial review and appointment of a substitute decision maker if incompetent, our data suggest that the number of patients could create enormous costs in money, effort, and delay. Judicial involvement undoubtedly will be required in some cases. But if policy seeks to assist the larger number of impaired patients that our study appears to identify, then it is likely that nonjudicial mechanisms should play a primary role in such efforts.

Clinical Implications

These data provide some guidance for clinicians who must make judgments about patients' competence, whether for purposes of initiating legal competence proceedings or exercising their authority to make the decision independently as provided by their state's laws.

First, all patient groups (as well as nonpatient groups) manifested considerably better understanding of the treatment information after it was disclosed to them part by part the second time (element disclosure) than when disclosed as a whole the first time (uninterrupted disclosure). This suggests that some patients who at first may seem to have deficiencies in their understanding of information may benefit by additional explanation. These data illustrate that a fixed inability to understand should not automatically be presumed on the basis of a patient's initial failure to comprehend the disclosure. Moreover, attempts to explain or "teach" the information may be of benefit for some patients, thereby avoiding unnecessary declarations of incompetence. On the other hand, the data also indicate that additional explanation is unlikely to improve understanding for a subgroup of severely disturbed patients.

Second, the results suggest that a diagnosis of schizophrenia should increase one's attention to the possibility of deficiencies in abilities related to legal competence, but that the diagnosis itself is only moderately related to serious deficits in those abilities. Those deficits are more likely to be substantial in cases involving greater degrees of thought disturbance (e.g., conceptual disorganization, active delusions, and hallucinations). Such cases, however, should be seen simply as "at greater risk" of incompetence. The relation is not strong enough to presume that serious thought disturbance "identifies" deficits in understanding, reasoning, or appreciation of a type sufficient to presume that the person is not able to make treatment decisions. Having identified a patient "at risk," additional assessment is needed in which the patient's actual ability to understand, reason about, and appreciate the significance of the disclosed information is evaluated in greater detail. Clinical strategies for performing such evaluations have been described elsewhere (e.g., Appelbaum & Grisso, 1988).

Where additional assessment is required in clinical cases, we would not recommend the use of the instruments that were developed for the present study. As noted earlier, research demands required that the instruments be more lengthy and complex than would be desirable for most clinical purposes. In addition, standardization of measurement required that the content of the UTD and TRAT be uniform across patients within a diagnostic category. This necessitated presentation of the treatment disclosure and the reasoning problem as hypothetical circumstances. In contrast, assessment of similar abilities in actual clinical cases should present information that represents the patient's actual disorder and clinical circumstances.

We are in the process of developing and validating an instrument designed specifically for standardized clinical assessments of abilities related to competence to consent to treatment. The MacArthur Competence Assessment Tool for Treatment Decisions (MacCAT-T) will use features of the research instruments employed in this study, and will assess abilities related to each of the four legal standards for competence. It will require only brief administration time, and will allow assessment of abilities in the context of the patient's own specific symptoms and treatment options.

REFERENCES

Appelbaum, P. S. (1988). The right to refuse treatment with antipsychotic medications: Retrospect and prospect. *American Journal of Psychiatry, 145,* 413–419.

Appelbaum, P. S., & Grisso, T. (1988). Assessing patients' capacities to consent to treatment. *New England Journal of Medicine, 319,* 1635–1638.

Appelbaum, P. S., & Grisso, T. (1992). *Manual: Perceptions of disorder.* Worcester, MA: University of Massachusetts Medical Center.

Appelbaum, P. S., & Grisso, T. (1995). The MacArthur Treatment Competence Study, I: Mental illness and competence to consent to treatment. *Law and Human Behavior, 19,* 105–126.

Appelbaum, P. S., & Roth, L. (1982). Treatment refusal in medical hospitals. In President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Making health care decisions,* Vol. 2. Washington, DC: U.S. Government Printing Office.

Appelbaum, P. S., & Schwartz, W. F. (1992). Minimizing the social cost of choosing treatment for the involuntarily hospitalized mentally ill patient: A new approach to defining the patient's role. *Connecticut Law Review, 24,* 433–485.

Beck, A. (1978). *Beck Depression Inventory.* San Antonio TX: Psychological Corporation.

Clare, L., McKenna, P., Mortimer, A., & Baddeley, A. (1993). Memory in schizophrenia: What is impaired and what is preserved? *Neuropsychologia, 31,* 1225–1241.

Gold, J., & Harvey, P. (1993). Cognitive deficits in schizophrenia. *Psychiatric Clinics of North America, 16,* 295–312.

Grisso, T., & Appelbaum, P. S. (1991). Mentally ill and nonmentally-ill patients' abilities to understand informed consent disclosures for medication. *Law and Human Behavior, 15,* 377–388.

Grisso, T., & Appelbaum, P. S. (1992). *Manual: Understanding treatment disclosures.* Worcester MA: University of Massachusetts Medical Center.

Grisso, T., & Appelbaum, P. S. (1993). *Manual: Thinking rationally about treatment.* Worcester MA: University of Massachusetts Medical Center.

Grisso, T., Appelbaum, P. S., Mulvey, E., & Fletcher, K. (1995). The MacArthur Treatment Competence Study, II: Measures of abilities related to competence to consent to treatment. *Law and Human Behavior, 19,* 127–148.

Guy, W. (Ed.) (1976). *ECDEU assessment manual for psychopharmacology (revised).* Rockville, MD: U.S. Department of Health, Education and Welfare.

Hedlund, J., & Vieweg, B. (1980). The Brief Psychiatric Rating Scale: A comprehensive review. *Journal of Operational Psychiatry, 11,* 48–65.

Hoge, S. K. (1993). Personal communication.

Hoge, S. K., Appelbaum, P. S., Lawlor, T., Beck, J. C., Litman, R., Greer, A., Gutheil, T. G., & Kaplan, E. (1990). A prospective, multicenter study of patients' refusal of antipsychotic medication. *Archives of General Psychiatry, 47,* 949–956.

Hollingshead, A., & Redlich, F. (1958). *Social class and mental illness.* New York: John Wiley.

Lidz, C., Meisel, A., Zerubavel, E., Carter, M., Sestak, R., & Roth, L. (1984). *Informed consent: A study of decisionmaking in psychiatry.* New York: Guilford.

Marcus, S. C., Robins, L. N., Bucholz, K. K., & Przybeck, T. R. (1989). *Diagnostic interview schedule screening instrument.* St. Louis, MO: Department of Psychiatry, Washington University School of Medicine.

Overall, J. (1988). The Brief Psychiatric Rating Scale (BPRS): Recent developments in ascertainment and scaling. *Psychopharmacology Bulletin, 24,* 97–99.

Overall, J., & Gorman, D. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports, 10,* 799–812.

Overall, J., & Klett, C. (1972). *Applied multivariate analysis.* New York: McGraw-Hill.

Rogers v. Commissioner, Dept' of Mental Health, 458 N.E. 2d 308 (Mass. 1983).

Wechsler, D. (1955). *The measurement and appraisal of adult intelligence.* Baltimore: Williams & Wilkins.

Zinermon v. Burch, 494 U.S. 113 (1990).