



# U.S. Department of Veterans Affairs

Public Access Author manuscript

*Ethics Law Aging Rev.* Author manuscript; available in PMC 2016 June 15.

Published in final edited form as:

*Ethics Law Aging Rev.* 2004 August 1; 10: .

## Hopes and Cautions for Instrument-Based Evaluation of Consent Capacity: Results of a Construct Validity Study of Three Instruments

**Jennifer Moye,**

Harvard Medical School

**Annin R. Azar,**

University of Louisville

**Michele J. Karel,** and

Harvard Medical School

**Ronald J. Gurrera**

Harvard Medical School

### Abstract

Does instrument based evaluation of consent capacity increase the precision and validity of competency assessment or does ostensible precision provide a false sense of confidence without in fact improving validity? In this paper we critically examine the evidence for construct validity of three instruments for measuring four functional abilities important in consent capacity: understanding, appreciation, reasoning, and expressing a choice. Instrument based assessment of these abilities is compared through investigation of a multi-trait multi-method matrix in 88 older adults with mild to moderate dementia. Results find variable support for validity. There appears to be strong evidence for good hetero-method validity for the measurement of understanding, mixed evidence for validity in the measurement of reasoning, and strong evidence for poor hetero-method validity for the concepts of appreciation and expressing a choice, although the latter is likely due to extreme range restrictions. The development of empirically based tools for use in capacity evaluation should ultimately enhance the reliability and validity of assessment, yet clearly more research is needed to define and measure the constructs of decisional capacity. We would also emphasize that instrument based assessment of capacity is only one part of a comprehensive evaluation of competency which includes consideration of diagnosis, psychiatric and/or cognitive symptomatology, risk involved in the situation, and individual and cultural differences.

### Introduction

#### Overview of Consent Capacity

Evaluations of capacity to consent to medical treatment are one form of competency evaluation. Questions about an individual's capacity to consent are typically raised when an individual with significant cognitive or psychiatric disorder is considerably impaired in her or his ability to understand and evaluate treatment information or to make and communicate treatment decisions. Often, when the issue involves the capacity to consent to ordinary treatment in the health care context, a clinician assesses the patient's decisional capacities. If

appropriate, a substitute decision-maker, such as a designated health care proxy, may make the current treatment decision. Evaluations of capacity to consent to medical treatment are rarely subject to judicial review (Grisso & Appelbaum, 1998). Yet, even in cases when assessments do not proceed to legal adjudication, legal definitions of capacity should inform clinical evaluations.

Questions about an individual's capacity to consent to treatment are raised most often when individuals disagree with what clinicians or family members believe to be a prudent course of treatment (Roe, Lereya, & Fennig, 2001). It is at these times that a clinician may evaluate the patient's clinical capacity to make an informed treatment choice. Because of the potential loss of self-determination, it is imperative that these evaluations are performed in a reliable and valid manner which accommodates individual differences and the specific situation.

### **Older Adults as Targets of Capacity Evaluation**

In the past, capacity evaluations targeted adults with serious and persistent psychiatric illness (e.g., severe schizophrenia) or neurological illness (e.g., neurodevelopmental disorder). More recently, as the population ages, capacity evaluations are applied to individuals with late onset conditions affecting cognition, such as dementia (Zimny & Grossberg, 1998). Such evaluations can present particular challenges. Adults with dementia may have a range of preserved or impaired abilities, depending on disease stage and etiology. As such, clinicians evaluating capacity in older adults may find themselves particularly challenged in balancing patient autonomy versus protection (Lo, 1990) in individuals with a life time of experience and perhaps subtle changes in cognitive abilities. Furthermore, unlike younger adults with earlier onset psychiatric or neurological conditions, older adults with late onset dementia may have considerable financial assets, at times introducing conflicts of interest for some parties.

### **Maximizing Patient Participation in Decision Making**

There are many ways to maximize an individual's participation in and capacity for health care decision making. Clinicians have the obligation to present information about the condition and treatments in clear, understandable language, accounting for difficulties in hearing or vision and the individual's educational level and language (Halter, 1999). Providing information in written or visual formats is helpful, as is addressing anxieties about conditions, treatments, and the management of possible consequences of these (e.g., pain). In some cases, it is useful simply to give the patient time to adjust to the impact of a new diagnosis and to consider treatment options (Grisso & Appelbaum, 1998).

### **Recognizing the Socio-Cultural Context of Decision Making**

Health care decision making most often occurs in a social context. Frequently, patients are supported by family members who together with the patient work to understand the condition and treatment options, and to make an appropriate treatment choice for the patient in light of her or his values. It is also important to be sensitive to cultural or spiritual beliefs affecting treatment choices, as well as the extent to which someone desires to be involved in health care decision making in the first place (Caralis, Davis, Wright, & Marcial, 1993).

## Problems in Clinical Evaluations of Consent Capacity

**Low Reliability of Clinical Evaluations**—Most evaluations of decisional capacity are made on the basis of a clinical interview. However, such clinical evaluations can be unreliable (Kaplan, Strange, & Ahmed, 1988; Markson, Kern, Annas, & Glantz, 1994; Marson, McInturff, Hawkins, Bartolucci, & Harell, 1997; Rutman & Silberfeld, 1992). In an important empirical investigation of the reliability of competency determinations for older adults, Marson and colleagues found that physicians achieved only a 56% judgment agreement ( $Kappa=.14$ ) in evaluating competency to consent in mildly demented patients (Marson, et al., 1997). Subjective impressions, experience in the field, or even ageism may influence a clinician's tolerance for allowing risk and respecting preferences of older adults (Clemens & Hayes, 1997).

**Challenges in Understanding Decision Outcomes**—In evaluating the extent to which another person's medical decision is rational, experts may have difficulty appreciating the elderly patient's perspective (Macklin, 1986). Clinicians may evaluate a patient's quality of life differently, typically as less desirable, than does the patient (Starr, Pearlmann, & Uhlmann, 1986; Uhlmann & Pearlmann, 1991). As it turns out, family and physician proxies are poor at predicting patients' treatment preferences (Seckler, Meier, Mulvihill, & Cammer Paris, 1991; Suhl, Simons, Reedy, & Garrick, 1994). Given individual differences in cultural background, religious beliefs, and life experiences, individuals may vary in the weighing of values that inform treatment decisions.

**Challenges in Understanding Decision Processes**—Cognizant of the nature of individual differences in treatment choices, clinicians who are evaluating consent capacity focus rightly on decisional processes not outcomes. However, evaluating decisional processes may become more challenging with age. As individuals grow older and gain more experience with life decisions, they appear to depend less on an analytic or rule-based decision process, and are more likely to rely upon automatic processes (Yates & Patalano, 1999). As such, it may be more difficult for the older adult to articulate and explain how she or he reached a particular decision. Furthermore, while older adults may be at an advantage or at least equivalent in tasks requiring automatic processing and implicit priming, they are at a distinct disadvantage on tasks requiring effortful processing and working memory (Craik & Jennings, 1992; Park, et al., 1996; Salthouse, 1996). These disadvantages would increase in situations where current illnesses, multiple medications, and situational anxieties are further taxing cognitive resources, such as would likely be the case when an individual's capacity to consent to treatment is being questioned.

## Legal Guidance in Evaluating Consent Capacity

**Statutory Definitions of Incapacity**—Clinicians may turn to the law for guidance in evaluating capacity. In the United States, most statutory definitions of incapacity include four parts: a) a disease or disorder, that causes; b) cognitive impairment, that leads to; c) functional (behavioral or decisional) deficits on a specific task or decision, that exceeds; d) an acceptable risk threshold to the person or society (Anderer, Coleman, Lichtenstein, & Barry, 1990; Sabatino & Basinger, 2000). The disease or disorder is the medical diagnosis that is the presumed cause of the cognitive and functional incapacity. Legal definitions of

incapacity have moved away from diagnosis-based definitions to “functional” definitions (Hommel, 1996). The emphasis on function means that the courts are looking for information about performance on the specific capacities in question. This functional information may refer to ADL’s if the capacity in question is independent living; it could be decisional abilities if the capacity in question is medical decision-making.

**Four Case Law Standards or Functional Abilities for Consent Capacity**—In evaluating the “functional” component of consent capacity, clinicians can focus on a framework of abilities relevant for legal competency to consent to treatment: (1) understanding, (2) appreciation, (3) reasoning, and (4) expressing a choice, as presented in Table 1. These abilities, or legal standards, were first identified by review of case law and statutes (Roth, Meisel, & Lidz, 1977) and refined further by legal scholars, clinicians, and ethicists (Berg, Appelbaum, & Grisso, 1996; Tepper & Elwork, 1984; see Grisso & Appelbaum, 1998 for more discussion). This legal framework is broadly similar to cognitive models that emphasize comprehending and encoding of treatment information, processing information and making a treatment decision, and communicating that treatment decision (Marson & Harrell, 1999; Moye, 1996).

### **The Problem of Linking Clinical Evaluations to Legal Standards**

Although the law theoretically provides some guidance in capacity assessment, it is often not clear how to translate legal definitions into clinical practice. While legal terms may seem at times vague, the process of legal argument by opposing parties is used as the tool by which the judge or jury applies the legal term in a specific case (Melton, Petrila, Pythress, & Slobogin, 1987). This sort of opposing argument protection is challenging to accomplish in clinical settings. Biomedical ethics panels may be the closest counterpart. However, more often, when consent capacity is in question, there is not the tool of argument and opposing parties. Instead, the “expert” opinion of a single evaluator is the deciding factor. The common use of a single evaluator to apply rather vague legal concepts is likely a chief contributor to the low reliability of these assessments in complex cases (Marson, et al., 1997; Moye, 2000).

### **Empirical Contributions for Evaluating Consent Capacity**

Several researchers have tried to improve the reliability of capacity evaluations by suggesting questions, interview guides, or more structured tools for assessing functional abilities in terms of legal standards (Edelstein, 2000; Janofsky, McCarthy, & Folstein, 1992; Grisso & Appelbaum, 1998; Marson et al., 1995). Such endeavors hold the promise of translating legal terms into clinically useful techniques. However, some worry that these efforts might eventually over-simplify the task, resulting in “capacimeters” that actually decrease reliability, validity, and rights’ protections (Kapp & Mossman, 1996). Indeed, those who have developed such instruments have also cautioned that they be used to contribute to, not substitute for, an individualized clinical evaluation and interpretation.

### **Summary**

In summary, we know that clinical judgments about consent capacity can be challenging and at times unreliable. This is especially true for older adults who have late onset neurological

conditions and subtle cognitive changes, who may also, because of age, use more automatic decisional processes. Fortunately, there has been an emerging consensus about the legal standards important for competency, and, new instruments that aim to assess these legal standards have been proposed. In our view, an important next step is the empirical study of the utility of these instruments. The long term goal of such research should be to support individualized yet reliable and valid assessments of decisional capacities, so that we can provide care that is maximally sensitive to our patients' preferences. Our bias is that there is a place for instruments in a comprehensive assessment of capacity. If used appropriately, these instruments should improve on more subjective evaluations by providing clinicians consistent means for operationalizing legal standards in clinical situations. Edelstein (2000) summarizes this middle-of-the-road position: "an approach residing somewhere between the use of unbridled clinical judgment and the use of a 'capacimeter' ... appears prudent." (p. 433).

## Method

### Research Goals

In this study, we examine and compare three instruments designed to assess capacity to consent to medical treatment. Each instrument provides specific questions and structure for the clinical evaluation of the four standards for legally competent decision making, that is, understanding, appreciation, reasoning, and choice. In this report we consider the evidence for construct validity of the instruments. The findings are reviewed in consideration of the operationalization of the four legal standards and implications for their measurement.

### Participants

Forty-four men and 44 women with dementia participated in the study. Mean age was 75.32 ( $\pm 6.23$ ) and mean years of education was 13.89 ( $\pm 3.00$ ). Years of age and education did not differ between the men and women. Three (3.4%) participants identified their race as African American, 1 (1.1%) identified his race as Native American and 84 (95.5%) identified their race as Caucasian. Ninety-eight percent of the participants reported English as their first language.

### Procedures

**Recruitment**—Participants were identified through fliers circulated in hospital waiting rooms, senior centers, senior housing, and advertisements placed in community newspapers, council of aging newsletters, and an Alzheimer Association newsletter. Efforts were made to recruit participants of color through special solicitation in organizations and media serving communities of color.

**Screening**—Interested adults telephoned the research coordinator, who completed a modified version of the Telephone Interview for Cognitive Status (TICS-M; Brandt, et al., 1993) assessing orientation, ten-word memory, naming, nonverbal praxis, attention and calculation. Because of the importance of delayed recall in the diagnosis of dementia, the TICS was modified to include delayed recall of the ten-word list for a total possible score of 50.

To exclude individuals with psychiatric conditions that might interfere with cognition, all participants completed the Geriatric Depression Scale - short form (GDS; Sheikh & Yesavage, 1986), and the Brief Symptom Inventory (BSI; Derogatis, 1975). Those individuals scoring above raw score of 10 on the GDS and a T score of 70 on the BSI were excluded.

To gather more information about dementia, all participants then completed the Dementia Diagnostic Screening Questionnaire (DDSQ; Rogers and Meyer, 1988), with the help of a caregiver if needed. The 94 item questionnaire asks about subjective memory difficulties, (e.g., “do you have trouble remembering things?”) and risk factors for specific dementia subtype etiologies (e.g., for vascular dementia, “have you ever had heart surgery?”; for substance induced persisting dementia, “have you ever stopped drinking because you felt you were drinking too much?”). Potential participants in the dementia group also provided medical records, including blood work, neurological examinations, and head CT or MRI. As this study did not focus on capacity impairments associated with specific forms of dementia, we included all individuals who met DSM-IV criteria for dementia attributable to any of the following subtypes: Alzheimer’s, Vascular, Parkinsons, Alcohol Induced persisting, or combined etiologies.

Clinical diagnoses of dementia were made by consensus of a geropsychiatrist and geropsychologist using DSM-IV criteria on the basis of cognitive screening scores; reports of memory and behavioral problems, presence of risk factors for dementia, and medical records. To avoid potential confounds with impairments due to delirium, all participants were medically stable at the time of testing. No participants were drawn from inpatient or long term care settings. In addition, all subjects had primary attention abilities within normal limits (WMS-III Digit Span combined score at least in the low average range; standard score 6).

**Informed Consent**—All participants completed informed consent as approved by hospital and medical school IRB and Human Subjects committees. As we aimed to recruit individuals in the early stages of dementia, we anticipated that most subjects would retain the capacities to consent to this low-risk study (not involving intervention, and with risks only of inconvenience, boredom, and fatigue), and would not have been adjudicated for legal incompetency. However, since our study involved individuals selected due to cognitive difficulties, special attention was given to the informed consent process. To complete informed consent, information about the study purpose, risks, and benefits was disclosed to all participants in simple direct language, in written and verbal formats. The examiner inquired as to whether the participant understood the study information, risks and benefits, and answered any questions. If the participant had a legal guardian (true for one subject in the study), the study information was disclosed conjointly to the participant and the guardian, and both signatures were obtained indicating the participant’s assent and the guardian’s consent. Participants were clearly informed that they could discontinue testing at any time if they disliked the testing or were fatigued. All but three participants finished the entire testing protocol. Subjects were compensated for their time.

**Testing**—Participants who screened into the study were invited to participate in face-to-face decision making capacity evaluation using standardized protocols. The order of the capacity instrument administration was counter balanced across the three instruments to avoid order effects. Testing occurred at the medical center, at a convenient location near the subject’s home (e.g., senior center) or in the subject’s home, depending upon the subject’s preference. Testing lasted approximately 120 minutes; participants were given scheduled breaks during testing.

Subjects were evaluated with three instruments designed to assess medical decision making capacity: (1) the MacArthur Competence Assessment Tool for Treatment (MacCAT-T; Grisso & Appelbaum, 1998); (2) the Hopemont Capacity Assessment Interview (HCAI; Edelstein, 2000); (3) the Capacity to Consent to Treatment Instrument (CCTI; Marson et al., 1995). Each instrument presents a hypothetical diagnosis and treatment alternatives, and asks the individual to demonstrate their capacity to understand and appreciate diagnostic and treatment information, and to explain their reasoning behind choosing one treatment alternative over another. Each instrument has a method for evaluating and rating the four legal standards, as presented in Table 2. In this study the MacCAT-T used a vignette involving treatment of a non-healing toe ulcer with surgery or amputation. The medical component of the HCAI has two vignettes about treatment of an eye infection and administration of cardio-pulmonary resuscitation in a friend; the use of the third person is intended to avoid personalization of information when used for research purposes (e.g., “but I don’t have that condition”). There is also a financial component to the HCAI not used in this study. The CCTI consists of brain cancer and heart surgery vignettes.

## Statistical Analyses

**Descriptive Statistics**—Descriptive data for each capacity variable were reviewed with special attention to distribution. Capacity instruments should have a low difficulty level, as healthy adults are presumed competent and should “test” so. In measurement terms, individuals without psychiatric or neurologic illness should score at or near ceiling on the measure, indicating adequate capacity. Individuals with significant psychiatric or neurologic illness should show some range of impairment on the capacity, if the capacity is likely to be impaired in that illness group. Individuals with mild to moderate degrees of illness would be expected to show mild to moderate impairment on relevant capacities, meaning some degree of left skew might be anticipated, although such non-normal distributions are not ideal for research purposes as they limit statistical analyses. Specifically, skewness and kurtosis values above 1.0 are undesirable.

**Multi-Trait Multi-Method Matrix**—Evidence for construct validity was evaluated through a multi-trait multi-method matrix, using four Campbell and Fiske (1959) criteria under Schmitt and Stults (1986) assumptions, in which each variable is hypothetically comprised of trait, method, and error variance. In this research, “trait” refers to the particular decisional capacity (understanding, appreciation, reasoning, expressing a choice) and “method” refers to the particular assessment tool (MacArthur, Hopemont, CCTI). Internal consistency reliability values (alpha) were computed for the reliability diagonals (mono-trait mono-method), when there was more than one item for the variable (shown in italics on Table 4).

Factor analytic and other multivariate methods of analyzing a multi-trait multi-method matrix could not be used due to restrictions in range and irregularities in distribution.

The Campbell and Fiske criteria for convergent or construct validity are:

1. Mono-trait hetero-method (MTHM) or “validity values” (shown in outline on Table 4) should be statistically significant and considerably different from zero.
2. MTHM validity diagonal values should be higher than hetero-trait hetero method (HTHM) (shown in bold on Table 4) values in the associated HTHM rows and columns; a variable should correlate higher with an independent method to measure the same trait than with independent methods to measure different traits.
3. MTHM validity diagonal values should be higher than hetero-trait mono-method (HTMM) values (shaded grey on Table 4) in the associated HTMM triangle; a variable should correlate higher with an independent method to measure the same trait than with the same method used to measure independent traits.
4. Patterns of trait interrelationships should be the same in mono-method and hetero-method triangles.

## Results

### Descriptive Statistics

Mean and standard deviation values for each capacity score are presented in Table 3, along with skewness and kurtosis values. For the capacity of understanding, the MacArthur has adequate range and within normal limits (WNL) distribution values, although there was very mild left skew. The Hopemont understanding mean score approaches the upper limit, and the skewness and kurtosis values are higher. The CCTI understanding mean score is much lower than the upper limit, but skewness and kurtosis are WNL, although there is mild left skew.

For the capacity of appreciation, the MacArthur mean is very near the upper limit with high skewness and kurtosis values. The Hopemont appreciation mean approaches the upper limit and the distribution is left skewed. CCTI appreciation range and distribution of scores are adequate.

For the capacity of reasoning, ranges and distributions are adequate on the MacArthur and CCTI, but the mean approaches the upper limit and skewness and kurtosis values are high for the Hopemont. For the capacity of choice, means on all three instruments approach the upper limit, and all distributions are considerably left skewed with high kurtosis values.

### Multi-trait Multi-method Matrix

The Multi-trait Multi-method matrix is shown in Table 4.

**Understanding Validity Values: Cambell-Fiske Criteria 1–3**—On the capacity of understanding, each of the MTHM validity values met the first three Campbell-Fiske



criteria. All three MTHM validity values were statistically significant and greater than zero. The MacArthur/Hopemont validity value (.77) was higher than the associated HTHM values, which ranged from  $r = .02$  to  $.57$  ( $M = .34$ ,  $SD = .16$ ) and HTMM values (range  $r = -.03$  to  $.74$ ;  $M = .36$ ,  $SD = .23$ ). The Hopemont/CCTI validity value (.74) was generally higher than the associated HTHM values, ranging from  $r = -.02$  to  $.56$  ( $M = .36$ ,  $SD = .19$ ) and HTMM values (range  $r = -.01$  to  $.74$ ;  $M = .35$ ,  $SD = .22$ ). The MacArthur/CCTI validity value (.74) was higher than the associated HTHM values, which ranged from  $r = -.02$  to  $.57$  ( $M = .28$ ,  $SD = .18$ ) and HTMM values (range  $r = -.03$  to  $.50$ ;  $M = .23$ ,  $SD = .18$ ).

**Appreciation Validity Values: Cambell-Fiske Criteria 1–3**—For the capacity of appreciation, the MTHM validity values ranged from  $.15$  to  $.20$ , none of which were statistically significant and all of which were in the range or lower than the reliability values ( $\alpha = .19$ ,  $.49$ ). In general, none of the values met the Campbell-Fiske criteria for the capacity of appreciation. The MacArthur/Hopemont validity value (.20) was lower than the mean associated HTHM values (range  $r = .00$  to  $.50$ ;  $M = .29$ ,  $SD = .16$ ) and HTMM values (range  $r = -.03$  to  $.74$ ;  $M = .36$ ,  $SD = .23$ ). The Hopemont/CCTI validity value (.19) was about the same as the mean score for the associated HTHM values (range  $r = -.04$  to  $.42$ ;  $M = .18$ ,  $SD = .14$ ) and lower than the HTMM mean score values (range  $r = -.01$  to  $.74$ ;  $M = .35$ ,  $SD = .22$ ). The MacArthur/CCTI validity value (.15) was lower than most of the associated HTHM values, which ranged from  $r = -.04$  to  $.50$  ( $M = .26$ ,  $SD = .16$ ) and HTMM values (range  $r = -.03$  to  $.50$ ;  $M = .23$ ,  $SD = .18$ ).

**Reasoning Validity Values: Cambell-Fiske Criteria 1–3**—For the capacity of reasoning, results were mixed. The MTHM validity values ranged from  $.26$  to  $.42$ , two of which were statistically significant. These values were somewhat smaller than reliability values which ranged from  $\alpha = .39$  to  $.59$ . Most of the Campbell-Fiske criteria were met for the reasoning validity values, although MTHM validity values were often close in size to mean scores for the relevant HTHM and HTMM correlations. The MacArthur/Hopemont MTHM reasoning validity value (.42) was higher than the mean scores for associated HTHM correlations, which ranged from  $r = .02$  to  $.57$  ( $M = .33$ ,  $SD = .18$ ) and HTMM correlations (range  $r = -.03$  to  $.74$ ;  $M = .36$ ,  $SD = .23$ ). The Hopemont/CCTI validity value (.40) was higher than the means of associated HTHM correlations, ranging from  $r = .01$  to  $.57$  ( $M = .28$ ,  $SD = .20$ ) and HTMM correlations (range  $r = -.01$  to  $.74$ ;  $M = .35$ ,  $SD = .22$ ). The MacArthur/CCTI validity value (.26) was lower than the mean relevant HTHM correlations (range  $r = .01$  to  $.47$ ;  $M = .30$ ,  $SD = .15$ ) and close to the mean of associated HTMM correlations values (range  $r = -.03$  to  $.50$ ;  $M = .23$ ,  $SD = .18$ ).

**Expressing a Choice Validity Values: Cambell-Fiske Criteria 1–3**—For the capacity of choice, results were mostly negative. The MTHM validity values ranged from  $r = -.07$  to  $.30$ , only one was statistically significant. Reliability values could be computed for two measures and were low ( $\alpha = .03$  and  $.32$ ). Most of the Campbell-Fiske criteria were not met for the capacity of expressing a choice. The MacArthur/Hopemont MTHM validity value ( $-.07$ ) was lower than most of the relevant HTHM correlations (range  $r = -.04$  to  $.56$ ;  $M = .19$ ,  $SD = .24$ ) and HTMM correlations (range  $r = -.03$  to  $.74$ ;  $M = .36$ ,  $SD = .23$ ). The Hopemont/CCTI validity value (.30) was lower than most associated HTHM

correlations (range  $r=.13$  to  $.56$ ;  $M=.32$ ,  $SD=.15$ ) and the HTMM values (range  $r=-.01$  to  $.74$ ;  $M=.35$ ,  $SD=.22$ ). The MacArthur/CCTI validity value (.25) was higher than most of the associated HTHM correlations (range  $r=-.04$  to  $.29$ ;  $M=.10$ ,  $SD=.12$ ) and close to the mean of the HTMM correlations values (range  $r=-.03$  to  $.50$ ;  $M=.23$ ,  $SD=.18$ ).

**Cambell-Fiske Criterion 4**—In comparing patterns of trait interrelations across three mono-method and six hetero-method triangles, correlations between measures of the capacities of understanding and reasoning were often the highest (8 of 9 triangles). Correlations between measures of appreciation and expressing a choice were many times lowest. Correlations of measures of reasoning and expressing a choice were often times mid rank. Correlations between appreciation and understanding, appreciation and reasoning, and expressing a choice and understanding varied in rank size across the triangles.

## Discussion

In this paper we consider the appropriateness of instrument based assessment of consent capacity. Instrument based assessment may address the low reliability of evaluations using interview alone and the difficulty of interpreting legal terms in clinical settings. To address the pros and cons of instrument based assessment we considered the evidence for construct validity associated with three instruments designed to evaluate four specific capacities in a group of 88 older adults with dementia.

Prior to the discussion of results, several limitations in regards to our sample are noted. First, we recruited individuals with dementia of heterogeneous causes. While we feel this increases the generalizability of our findings to the types of patients seen in our clinical practice, who most often present with multiple risk factors for different forms of dementia, it means we cannot comment on patterns of deficits associated with specific forms of dementia (e.g., Alzheimer's). Demographic variables suggest our sample is not equivalent to the United States population. Although we made special efforts to recruit individuals of color, 96% of our participants with dementia were Caucasian. Also, 39% of our sample had a high school education or less (lower than the current national mean within the 60+ age group). In addition, although our income data are comparable to the national median (within the 60+ age group), financial compensation may have attracted individuals with greater financial needs. Forty-three percent of our participants with dementia were recruited from V.A. clinics, which likely influenced these sample characteristics.

## Understanding

The capacity of understanding concerns whether the individual can demonstrate comprehension of basic diagnostic and treatment information. The CCTI and Hopemont assess understanding by asking for recall of specific facts about the diagnostic and treatment information. The MacArthur assesses understanding by asking for paraphrasing of information in a series of disclosure and cued recall prompts.

Distribution results suggest that there is some tendency to left skew (majority of scores approaching ceiling), which seems appropriate for individuals with mild dementia; this is to say that capacity instrument scores would not always show normal distributions. Incapacity

should be found only with significant impairment, thus mildly impaired individual's should score near ceiling. In this study the left skew was especially true for the Hopemont, designed to be appropriate for more considerably impaired individuals within long term care settings.

Examination of the multi-trait multi-method matrix showed good convergent validity for the capacity of understanding as measured by the three instruments, with high inter-test correlations, which generally exceeded correlations between different traits by different measures and by the same measures. These findings suggest good agreement between instruments for the measurement of understanding, and support the concept of understanding and its measurement through structured instruments.

### Appreciation

The capacity of appreciation is somewhat complex. In case law it has been applied to individuals who may show superficial understanding of and reasoning about medical information, yet fail to accept that they have the condition or that the treatment may be of benefit. As such it relates to trust and judgment and applies especially to the interpersonal aspects of medical decisions (whether a patient believes a doctor). Problems with appreciation may be seen in individuals with delusional disorders, but could occur in any individual who struggles with discernment and belief.

The three instruments used here assess it differently: the MacArthur as “reasons to doubt” the disorder or treatment benefit – on which scores are considerably left skewed; the Hopemont, as direct questions about why the doctor is suggesting treatment – for which scores are somewhat left skewed; the CCTI, as future planning and projection in the context of planned treatments – on which distributions are more normal. These distributions suggests that individuals with dementia may not show appreciation impairments through doubting of information, or questioning doctors. As such, the CCTI approach of assessing appreciation is rather creative and may be useful for individuals with dementia.

It is not surprising, given the different methods of measuring this complex construct and the limited observed variance, that the evidence for convergent validity for the concept of appreciation was poor. While the CCTI approach for measuring appreciation in demented populations may hold promise, more research is needed to establish the meaning of the concept in different populations and valid means for its assessment. Such research should compare different ways of measuring appreciation and explore the potential roles of social judgment (e.g., trust of clinician) and reflective judgment or insight, foresight, and discernment (e.g., trust of information and personal reflection upon it).

### Reasoning

The capacity of reasoning involves the ability to weigh risks and benefits of treatment alternatives, and also to compare the impact of those risks and benefits on one's own life. The outcome of this reasoning is the treatment choice. To evaluate an individual's capacity to reason, the examiner asks the patient to back track and articulate the internal and perhaps intuitive process of arriving at that choice. The assessment of reasoning is operationalized on the CCTI and Hopemont as the provision of “rational reasons,” in many ways a restatement of the facts previously recounted for the assessment of understanding, as supported by the

high correlations between understanding and reasoning as measured by the CCTI and Hopemont. On the MacArthur, reasoning is assessed by comparing treatments in light of stated consequences (e.g., immobility), and how those affect the individual's life and valued activities (e.g., inability to participate in certain activities). The first part of the MacArthur assessment of reasoning is similar to the CCTI and Hopemont, while the second part, involving application to one's own life, extends the task.

The evidence for convergent validity was mixed, although generally positive. The association between the MacArthur and CCTI methods of assessing reasoning had the lowest correlation, and was often lower than the correlations of different traits by different methods. Other hetero-method correlations for reasoning were higher and met Campbell-Fiske criteria. These findings suggest mixed agreement between instruments for the measurement of reasoning. More research is needed to compare methods that focus on the capacity to list rational reasons versus the ability to articulate how risks and benefits mesh with one's own interests and values. We favor the latter approach, as it seems the process of asking the individual to explain the reasoning process in light of values (Karel, 2000) would seem a more rich assessment of reasoning than simply the re-stating of reasons (however much that may comply with some views of case law). In our view, "competent" reasoning may be best described as a personally authentic or valid decision process rather than a logical or rational one (Snyder, 1994).

## Choice

The capacity of expressing a choice was evaluated as the ability to state a clear enough choice for oneself (MacArthur, CCTI) or another (Hopemont). In this sample, almost all individuals could do this, as indicated by highly left skewed distributions. The capacity of expressing a choice may be most relevant to individuals in more advanced stages of dementia, for whom the simple ability to state a choice may be impaired. In those cases it is probably not necessary (or possible) to assess other capacities. For individuals in early stages of dementia, such as in this sample, the simple capacity to express a choice is less likely to be impaired. We think the concept has meaning in general, but its convergent validity in this sample was limited by restrictions in range.

## Pros and Cons of Instrument Based Assessment of Consent Capacity

Is the move to develop instruments to evaluate legal competency advantageous? Do instruments increase precision of measurement so that the specific capacity is more reliably and meaningfully measured; or do instruments in their sense of precision merely prevaricate by confusing precision with validity? Of note, the answer to this question should be oriented around the type of capacity in question. An instrument to assist in the evaluation of capacity to drive or capacity to live at home would be quite different from one that assesses capacity to consent to medical treatment.

As regards instruments for consent capacity, the findings of this study suggest that there is relative agreement on the assessment of the capacity of understanding diagnostic and treatment information, and probably (although our data were too restricted in range to be conclusive) on the assessment of expressing a choice (a threshold ability). The concepts of

reasoning, and especially appreciation, need more fleshing out before we can feel confident that they are being assessed in a way that is meaningful across measures and for specific populations.

We suggest that it would be helpful to develop trial instruments that draw from the strengths of each of these instruments and also explore novel means to assess specific capacities. Especially for the capacities of reasoning and appreciation, it would be interesting to study multiple items or means for measuring the capacities within a single instrument, so that these items can be compared within a single test or method through item-scale or factor analyses. The best items or methods for assessing the capacity can then be retained. Also, although it can be hard to fund basic psychometric research, it is imperative to establish test re-test reliability and normative properties of these instruments.

In our caution about avoiding a rush to develop and use capacity instruments in clinical work, we are nevertheless mindful of the limitations of contemporary practice without such instruments. Subjective clinical interviews, of questionable reliability, are still the norm in evaluation of consent capacity. In long term care facilities, assessing capacity to complete an advance directive most often relies on only the evaluation of orientation to person, place, or time as an indication of that capacity; such capacity assessments rarely involve physicians and sometimes rely on clerical personnel (Walker, Blechner, Gruman, & Bradley, 1998). Against this norm, instruments that direct evaluation towards consideration of appropriate cognitive tasks and legal standards are certainly worth further consideration and development.

In closing, we return to a critical point illustrated. Capacity instruments are meant to practically and directly assess performance on a specific task or decision in question. But this component does not alone comprise a capacity evaluation, because it does not relate that performance to general abilities nor the diagnosis that explains why, on what basis, and for how long the observed capacity impairments may exist. Nor does it relate the capacity impairments to individual, situational, familial, and cultural factors of the context. If we can continue to improve capacity instruments and remember to place these within the framework of a comprehensive evaluation, we should ultimately be able to improve the reliability and validity of these important assessments.

## Acknowledgments

This research was supported by a NIMH grant to the first author (R29 MH57104). We thank Barry Edelstein, Jayne Fulkerson, Linda Ganzini, Daniel Marson, and Stacey Wood for helpful comments on earlier drafts of this paper. Special gratitude is extended to Thomas Grisso whose wise counsel from the outset of this project exemplifies gracious mentorship of new investigators.

## References

- Anderer, S.J.; Coleman, N.A.; Lichtenstein, E.C.; Parry, J.W. Determining competency in guardianship proceedings. Washington D.C.: American Bar Association; 1990.
- Berg JW, Appelbaum PS, Grisso T. Constructing competence: Formulating standards of legal competence to make medical decisions. *Rutgers Law Review*. 1996; 48:345–396. [PubMed: 16086484]

- Brandt J, Welsh KA, Breitner JC, Folstein MF, Helms M, Christian JC. Hereditary influences on cognitive functioning in older men: A study of 4000 twin pairs. *Archives of Neurology*. 1993; 50:599–603. [PubMed: 8503796]
- Campbell D, Fiske D. Convergent and discriminant validation by the multi-trait multi-method matrix. *Psychological Bulletin*. 1959; 54:81–105. [PubMed: 13634291]
- Caralis PV, Davis B, Wright K, Marcial E. The influence of ethnicity and race on attitudes toward advance directives, life-prolonging treatments, and euthanasia. *The Journal of Clinical Ethics*. 1993; 4:155–165. [PubMed: 8334279]
- Clemens E, Hayes HE. Assessing and balancing elder risk, safety, and autonomy: Decision making practices of elder care workers. *Home Health Care Services Quarterly*. 1997; 16:3–20. [PubMed: 10173442]
- Derogatis, LR. *Brief Symptoms Inventory*. Minneapolis: National Computer Systems; 1993.
- Edelstein B. Challenges in the assessment of decision-making capacity. *Journal of Aging Studies*. 2000; 14:423–437.
- Grisso, T.; Appelbaum, PS. *Assessing competence to consent to treatment*. New York: Oxford; 1998.
- Halter, JB. The challenge of communicating health information to elderly patients: A view from geriatric medicine. In: Park, DC.; Morrell, RW.; Shifren, K., editors. *Processing of medical information in aging patients: Cognition and human factors perspectives*. Mahwah, New Jersey: Lawrence Erlbaum; 1999. p. 23-28.
- Hommel, PA. Guardianship reform in the 1980s: A decade of substantive and procedural change. In: Smyer, M.; Schaie, KW.; Kapp, MB., editors. *Older adults decision making and the law*. New York: Springer; 1996. p. 225-253.
- Janofsky JS, McCarthy RJ, Folstein MF. The Hopkins Competency Assessment Test: A brief method for evaluating patients' capacity to give informed consent. *Hospital and Community Psychiatry*. 1992; 43:132–136. [PubMed: 1572608]
- Kaplan KH, Strange JP, Ahmed I. Dementia, mental retardation, and competency to make decisions. *General Hospital Psychiatry*. 1988; 10:385–388. [PubMed: 3203880]
- Kapp M, Mossman D. Measuring decisional capacity: Cautions on the construction of a "capacimeter". *Psychology, Public Policy, and Law*. 1996; 2:73–95.
- Karel MJ. The assessment of values in medical decision making. *Journal of Aging Studies*. 2000; 14:403–422.
- Lo B. Assessing decision making capacity. *Law, Medicine, and Health Care*. 1990; 18:193–201.
- Macklin, R. The geriatric patient: Ethical issues in care and treatment. In: Mappes, T.; Zambaty, JS., editors. *Biomedical Ethics*. 2nd. New York: McGraw-Hill; 1986. p. 162-167.
- Markson LJ, Kern DC, Annas GJ, Glantz LH. Physician assessment of patient competence. *Journal of the American Geriatrics Society*. 1994; 42:1074–1080. [PubMed: 7930332]
- Marson, DC.; Harrell, LE. Neurocognitive changes associated with loss of capacity to consent to medical treatment in patients with Alzheimer's disease. In: Park, DC.; Morrell, REW.; Shifren, K., editors. *Processing of medical information in aging patients: Cognition and human factors perspectives*. Mahwah, New Jersey: Lawrence Erlbaum; 1999. p. 109-126.
- Marson DC, Ingram KK, Cody HA, Harrell LE. Assessing the competency of patients with Alzheimer's disease under different legal standards. *Archives of Neurology*. 1995; 52:949–954. [PubMed: 7575221]
- Marson DC, McInturff B, Hawkins L, Bartolucci A, Harrell LE. Consistency of physician judgments of capacity to consent in mild Alzheimer's disease. *Journal of the American Geriatric Society*. 1997; 45:453–457.
- Melton, GB.; Petrila, J.; Poythress, NG.; Slobogin, C. *Psychological evaluations for the courts*. New York: Guilford; 1987.
- Moye J. Theoretical frameworks for competency assessments in cognitively impaired elderly. *Journal of Aging Studies*. 1996; 10:27–42.
- Moye J. Mr. Franks refuses surgery: Cognition and values in competency determination in complex cases. *Journal of Aging Studies*. 2000; 14:385–401.

- Park DC, Smith AD, Lautenschlager G, Earles J, Frieske D, Zwahr M, Gaines C. Mediators of long-term memory performance across the life span. *Psychology and Aging*. 1996; 11:621–637. [PubMed: 9000294]
- Roe D, Lereya J, Fennig S. Comparing patients' and staff members' attitudes: Does patients' competence to disagree mean they are not competent? *Journal of Nervous and Mental Disease*. 2001; 189:307–310. [PubMed: 11379974]
- Rogers RL, Meyer JS. Computerized history and self-assessment questionnaire for diagnostic screening among patients with dementia. *Journal of the American Geriatrics Society*. 1988; 36:13–21. [PubMed: 2961795]
- Roth LH, Meisel CA, Lidz CA. Tests of competency to consent to treatment. *American Journal of Psychiatry*. 1977; 134:279–284. [PubMed: 842704]
- Rutman D, Silberfeld M. A preliminary report on the discrepancy between clinical and test evaluations of competency. *Canadian Journal of Psychiatry*. 1992; 37:634–639. [PubMed: 1477822]
- Sabatino CP, Basinger SL. Competency: Reforming our legal fictions. *Journal of Mental Health and Aging*. 2000; 6:119–144.
- Salthouse TA. The processing-speed theory of adult age differences in cognition. *Psychological Review*. 1996; 103:403–428. [PubMed: 8759042]
- Schmitt N, Stults DM. Methodology review: Analysis of multi-trait multi-method matrices. *Applied Psychological Measurement*. 1986; 10:1–22.
- Seckler AB, Meier DE, Mulvihill M, Cammer Paris BE. Substituted judgement: How accurate are proxy predictions? *Annals of Internal Medicine*. 1991; 115:92–98. [PubMed: 2058868]
- Sheikh, JI.; Yesavage, JA. Geriatric depression scale: Recent evidence and development of a shorter version. In: Brink, TL., editor. *Clinical gerontology: A guide to assessment and intervention*. New York: Haworth; 1986.
- Snyder AC. Competency to refuse lifesaving treatment: Valuing the nonlogical aspects of a person's decisions. *Issues in Law and Medicine*. 1994; 10:299–320. [PubMed: 7829306]
- Starr TJ, Pearlman RA, Uhlmann RF. Quality of life and resuscitation decisions elderly patients. *Journal of General Internal Medicine*. 1986; 1:373–379. [PubMed: 3794836]
- Suhl J, Simons P, Reedy T, Garrick T. Myth of substituted judgement: Surrogate decision making regarding life support is unreliable. *Archives of Internal Medicine*. 1994; 150:90–96. [PubMed: 8267493]
- Tepper A, Elwork A. Competency to consent to treatment as a psychological construct. *Law and Human Behavior*. 1984; 8:205–223. [PubMed: 11659145]
- Uhlmann RF, Pearlman RA. Perceived quality of life and preferences for life-sustaining treatment in older adults. *Archives of Internal Medicine*. 1991; 151:495–497. [PubMed: 2001131]
- Walker L, Blechner B, Gruman C, Bradley E. Assessment of capacity to discuss advance care planning in nursing homes. *Journal of the American Geriatrics Society*. 1998; 46:1055–1056. [PubMed: 9706904]
- Yates, JF.; Patalano, AL. Decision making and aging. In: Park, DC.; Morrell, RW.; Shifren, K., editors. *Processing of medical information in aging patients: Cognition and human factors perspectives*. Mahwah, New Jersey: Lawrence Erlbaum; 1999. p. 31-54.
- Zimny, GH.; Grossberg, GT. Guardianship of the Elderly. In: Zimny, GH.; Grossberg, GT., editors. *Guardianship of the elderly*. New York: Springer; 1998. p. 3-24.

**Table 1**

## Definitions of Decisional Capacities Important in Legally Competent Decision Making

Understanding	The ability to comprehend diagnostic and treatment related information, and to demonstrate that comprehension; involves ability to attend, encode, store, and retrieve newly presented words and phrases
Appreciation	The ability to determine the significance of treatment information relative to one's own situation, focusing on beliefs about the actual presence of the diagnosis and the possibility that treatment would be beneficial; involves insight, judgment, and foresight
Reasoning	The process of comparing alternatives in light of consequences, through integrating, analyzing, and manipulating information; involves the ability to provide rational reasons for a treatment decision, to manipulate information rationally, to generate consequences of treatments for one's life, to compare those consequences in light of one's values
Expressing A Choice	This ability to communicate a decision about treatment, applying to individuals who cannot or will not express a choice, or who are ambivalent



**Table 2**

## Comparison of Methods of Assessing Capacities for each Instrument

Capacity	Method
Understanding	
MacCAT	Paraphrasing, “in your own words” of diagnostic and treatment information.
HCAI	Questions about information described in the condition and treatment.
CCTI	Questions about the details of information described in two vignettes.
Appreciation	
MacCAT	Asks (i) if there is “any reason to doubt” information about the condition; (ii) whether treatment “might be of benefit”.
HCAI	Asks why the doctor wants the person to take the treatment (in 1 <sup>st</sup> vignette).
CCTI	Asks (i) preparation: what would need to be done to prepare for the chosen treatment; (a) projection: what life will be like one year after the treatment.
Reasoning	
MacCAT	Asks for (i) comparison of treatments; (ii) consequences of treatments; (iii) everyday impact of treatment alternatives; (iv) logical consistency.
HCAI	Asks why a choice made, with query and credit for the risks and benefits considered.
CCTI	Asks to give all the reasons why made choice, credit given for the total number and accuracy of reasons provided.
Expressing A Choice	
MacCAT	Rates if there is a clear choice.
HCAI	Rates whether a choice is made.
CCTI	Rates whether a choice is made for each vignette.

Table 3

## Mean Values for Decisional Capacities

Capacity/Instrument	# Items	Potential Range	M	(SD)	Skewness	Kurtosis
Understanding						
MacCAT-T	12	0-24	16.92	(4.17)	-0.71	0.16
HCAI	14	0-28	24.19	(3.38)	-1.85	3.34
CCTI	18	0-130	45.99	(13.98)	-0.55	-0.17
Appreciation						
MacCAT-T	2	0-4	3.79	(0.54)	-2.98	10.04
HCAI	1	0-2	1.66	(0.52)	-1.15	0.29
CCTI	4	0-8	4.98	(1.60)	-0.14	-0.57
Reasoning						
MacCAT-T	4	0-8	6.67	(1.32)	-0.76	-0.34
HCAI	3	0-6	5.07	(1.29)	-1.71	3.02
CCTI	2	0-20 <sup>a</sup>	3.91	(1.92)	0.13	-0.60
Expressing A Choice						
MacCAT-T	1	0-2	1.96	(0.19)	-5.13	24.88
HCAI	3	0-6	5.79	(0.56)	-2.56	5.29
CCTI	2	0-4	3.94	(0.32)	-3.08	8.82

<sup>a</sup>CCTI total scores for reasoning were standardized to ten points per vignette to adjust for different potential total scores on reasoning associated with different treatment choices.

**Table 4**

Multi-trait (Capacity) Multi-method (Instrument) Matrix

		MacCAT				HCAI				CCTI			
		U	A	R	C	U	A	R	C	U	A	R	C
MacCAT	U	.77											
	A	.42*	.19										
	R	.48*	.34*	.42									
	C	-.03	.04	.24	+								
HCAI	U	.77*	.47*	.47*	.02	.74							
	A	.42*	.20	.31*	.00	.42*	+						
	R	.57*	.36	.42*	.02	.74*	.23	.59					
	C	.51*	.50*	.43*	-.07	.64*	.27	.53*	.03				
CCTI	U	.74*	.39*	.46*	-.02	.74*	.37*	.56*	.56*	.87			
	A	.24	.15	.15	-.04	.23	.19	.13	.18	.26	.49		
	R	.31*	.35*	.26	.01	.39*	.04	.40*	.37*	.50*	.28*	.39	
	C	.13	.19	.27	.25	.29*	.14	.21	.30*	.14	-.01	.20	.32

Note. U=Understanding; A=Appreciation; R=Reasoning; C=Expressing a Choice. Mono-trait Hetero-method (MTHM) validity values are outlined. Hetero-trait Mono-method (HTMM) values are shaded. Hetero-trait Hetero-method (HTHM) values are in bold. Mono-trait Mono-method reliability values are in italics.

\* Correlation is significant at the 0.01 level (2-tailed).

+ Alpha cannot be computed as there is only one item.