

# Closing the Loop: Assessing the Effectiveness of Psychiatric Competency Measures

David Bienenfeld, M.D.

William Klykylo, M.D.

Douglas Lehrer, M.D.

**Background:** During the past several years there has been a growing emphasis on competency assessment in graduate education in all specialties, including psychiatry. Methods of assessment are continuously being devised and refined. **Objective:** To outline principles and methods of assessing the validity of competency measures and their relevance to trainees. **Methods:** The authors reviewed the literature relevant to competency assessment and quality improvement in the fields of elementary and secondary education, adult learning, undergraduate medical education, and postgraduate medical training. Three means were used to evaluate the assessment program: 1) The authors surveyed graduates and credentialing agencies for feedback on the relevance of the measures they previously developed for current and envisioned future practice of psychiatry; 2) They measured the completion rates of competency measures by residents in the first 3 years of a program, measuring general psychiatric competencies; and 3) They surveyed residents and faculty regarding satisfaction with the existing process. **Results:** Despite small numbers, the results indicated that measures devised in 1998 were relevant to the practice of psychiatry in 2001 and envisioned by practitioners for the subsequent 5 to 10 years. Resident completion rates of competency requirements indicated that the structure of the measures was appropriately paced to parallel the learning expectations of the training program. Participant satisfaction with the process of developing and implementing the measures was high. **Conclusions:** While devising and implementing competency assessments is critical, the process should not stop with the application of the measures. Ongoing evaluation of these measures for their continued suitability is essential. Training programs can readily determine the effectiveness of competency programs and improve them as necessary. (*Academic Psychiatry* 2003; 27:—)

Much of the early work on competency assessment in postgraduate medicine has focused on the development of measures. Once a program has applied measurement (the comparison of individuals against some standard), it should follow with an evaluation of the effectiveness of the measurements. Because those who design the assessment plan are usually in charge of its implementation, it is often beneficial to recruit evaluators from outside and among users of the measures and afford these individuals the freedom to question the measures and objectives that support them (1).

In designing performance-based measures, it is easy to lose the forest for the trees (i.e., to keep breaking down skills into ever smaller steps that lose relevance). The best measures are designed with the ultimate life needs of the learner in mind as well as the reasons for the skills being measured (2,3). Evaluation of the competency program should therefore include some determination of how the skills being measured reflect the ultimate needs of the graduates. In most cases, best results are obtained when multiple sources are used (1,4). One dimension of program evaluation is the continued monitoring of the process. If the pro-

gram includes built-in measures of its function, teachers can monitor those measures and use them to modify and improve the program itself (1,5-7).

Additionally, teachers can evaluate the validity of their program. Validity is comprised of four types (2,8,9):

1. *Construct validity* reflects the degree to which the test results measure the student's ability to perform the required task;
2. *Content validity* reflects the degree to which completion of the required task reflects the objective of the program;
3. *Predictive validity* is the ability of the test to predict the likelihood that the student will be able to perform a task in the future; and
4. *Concurrent validity* is the degree to which a test measures skills possessed by those who have a degree of expertise that is not itself directly measurable.

In a previous publication, our group described the process by which we developed a measure of competency in a wide range of psychiatric skills (10). We emphasized the participatory nature of devising our schema and sought input from recent graduates as well as other community practitioners in selecting our objectives. We also included a preliminary report on the outcomes of the competency assessments within our training program after 1 year.

The competency measures described in our earlier report address skills across five content areas: assessment and presentation, diagnosis, somatic treatment, consultation liaison and medical psychotherapy, and psychotherapy. In each year, a resident is required to demonstrate the acquisition of specific skills a certain number of times. For example, in postgraduate year (PGY-I), a resident must correctly diagnose substance abuse eight times and demonstrate the safe and effective use of selective serotonin reuptake inhibitors (SSRIs) five times. In PGY-III, a resident must demonstrate safe and effective management of the sexual side effects of antidepressant or antipsychotic medications five times and provide psychotherapy for patients with dysthymic disorder or minor depression six times. Certification is provided by a supervisor's initials on the resident's log for each episode of demonstrated skill

acquisition (10). (See Figure 1 for a sample page of the checklist.)

The competency assessment program has been in place for more than 3 full years. Here, we here describe our evaluation of the program, including measures of ongoing performance and concurrent validity. Although this study evaluated a measure of broad psychiatry skills, it is included in this article as a generalized model for ongoing evaluation of competency measures.

---

## Methods

### Ongoing evaluation

Twice yearly, the residency training director (DB) meets with residents to review progress towards completion of competency objectives and documents the results. Records of these semiannual reviews were examined for all residents who graduated from 1999 to 2001. Records were rated as "complete" (end of year) or "on track to complete" (mid-year). For 1999 graduates, only PGY-IV reviews were examined. For

FIGURE 1. Sample pages of Skills Competency Checklist

<u>By the end of PGY-III</u>				
I. ASSESSMENT AND PRESENTATION				
1. Present a complete psychodynamic case formulation.				
II. DIAGNOSIS				
1. Correctly make the diagnosis of dissociative or delusional disorder.				
2. Ascertain the diagnostic criteria for mental retardation.				
3. Correctly make the diagnosis of attention deficit hyperactivity disorder.				
4. Correctly make the diagnosis of childhood depression.				

all other classes, PGY-III and PGY-IV records were studied.

### Resident and faculty evaluation

Surveys were sent to all current residents in the program and 15 members of the faculty who regularly certified residents' completion of behavioral objectives. Surveys addressed whether the goals of the training program were met, whether they correlated with competencies that were measured, and whether residents received feedback on their attainment of competencies. (Questions are reproduced with results in Table 1.)

### Evaluation by graduates and credentialing bodies

Surveys were sent to all graduates of the program (for whom current addresses could be obtained) from the 5 years before implementation of the competency assessment program. The same survey was sent to the medical directors of the eight members of the American Managed Behavioral Healthcare Association and the directors of psychiatric services for the 16 hospitals that were rated by *U.S. News* as providing the best psychiatric services (11). All recipients of the survey received a complete copy of the skills competency checklist.

## Results

### Ongoing evaluation

Table 1 presents the results of mid-year and end-of-year completion rates for competency requirements. At mid-year, approximately one-half of the residents were deficient, but deficiencies were rare by

TABLE 1. Competency checklist results at semiannual reviews

	Number of residents complete at year end or on-track to complete at mid-year*	Number of residents deficient
End of PGY-IV	13	1
Mid-PGY-IV	7	5
End of PGY-III	9	1
Mid-PGY-III	4	5

the end of each year. Of the mid-year deficiencies, eight of the 10 total deficiencies were in obtaining faculty initials for the completed tasks. The remaining two were the result of difficulty in obtaining cases. The single end-of-year deficiency in PGY-III was compensated for by mid-year of PGY-IV. The single end-of-residency deficiency was for a resident who had not completed therapy of a marital or family case, which was noted in his record. (Number of mid-year and end-year evaluations does not always match because of occasional missed mid-year evaluations.)

Of the mid-year deficiencies, most were in 1999, which was the first full year of implementation. By 2001, mid-year deficiencies, even in obtaining initials, were unusual (two PGY-III and one PGY-IV).

### Resident and faculty evaluation

Response rate was 19/30 (63%) for residents and 7/15 (47%) for faculty. (See Table 2 for results.) Faculty generally found implementation of the checklist uncomplicated. Most believed implementation of the checklist had a positive impact on clarifying objectives and the skills required for the checklist were consistent with the educational objectives of the program. They found little or moderate effect on their feedback to residents as a result of the competency requirements.

Residents were mostly positive about the ease of implementation—though not as enthusiastic as faculty—with few opinions regarding the procedure as moderately difficult. Most agreed with faculty that the checklist was useful in clarifying training objectives. Residents did not perceive much impact on the feedback they received from faculty. Most believed that the skills were moderately reflective of the training program's objectives.

### Evaluation by graduates and credentialing bodies

The response rate for graduates was 11/20 (55%) and 4/24 (17%) for credentialing bodies. Since response numbers were so small, both groups were combined. All respondents believed that the requirements reflected current and anticipated needs for practitioners' skills at least moderately well (see Table 3).

None of the skills were perceived as either un-

necessary or overemphasized. Respondents advocated that measures should place more emphasis on:

- financial and legal issues (8)
- administrative and leadership skills (4)
- substance abuse (3)
- electroconvulsive therapy (ECT)
- cognitive therapy
- psychoanalytic psychotherapy
- geriatric psychiatry
- community mental health

**Discussion**

The ongoing evaluation of our competency assessment program identified areas that individual residents needed to improve. The fact that there were a moderate number of mid-year deficiencies but only rare end-year deficiencies implies that the criteria are attainable and elaborative enough to measure formative competency. In the earliest cohorts, such deficiencies were more common than during succeeding years, indicating that greater familiarity with the

expectations was reflected in more consistent completion of the target skills.

This result implies that awareness of the objectives and their attendant requirements increased with time. The survey that was conducted 3 years after implementation confirmed that residents believed that the measurement of competency clarified the objectives.

We selected the strategy of examining concurrent validity to evaluate the correlation of our competency program with societal needs. Although response rates from graduates and representatives of credentialing bodies were small, the responses received were fairly consistent. Our skills requirements for the training program reflected the skills that practitioners and healthcare agencies believe are necessary for current and anticipated practice in the near future.

There are a number of shortcomings in our evaluation however. First, the use of completion rates as an evaluation measure may be tautological since the training director who performs the evaluation is also responsible for implementing corrections. Similarly,

**TABLE 2. Resident and faculty evaluation of the skills competency checklist program**

How easy or difficult has it been to determine what is expected for each of the requirements?					
	Very easy	Moderately easy	Neutral	Moderately difficult	Very difficult
Faculty	1	5	0	0	0
Residents	4	10	4	1	0
How easy or difficult has it been to provide (faculty)/obtain (residents) patients or procedures for completed requirements?					
	Very easy	Moderately easy	Neutral	Moderately difficult	Very difficult
Faculty	1	4	2	0	0
Residents	3	9	5	2	0
How easy or difficult has it been to provide (faculty)/obtain (residents) initials for completed requirements?					
	Very easy	Moderately easy	Neutral	Moderately difficult	Very difficult
Faculty	4	2	0	0	0
Residents	5	7	6	1	0
What influence, if any, have the competency requirements had on clarifying training objectives?					
	Very positive	Moderately positive	None	Moderately negative	Very negative
Faculty	1	5	1	0	0
Residents	3	12	4	0	0
To what extent have the competency criteria affected the performance feedback you have given (faculty)/received (residents)?					
	Very much	Moderately	Little	None at all	
Faculty	0	4	3	0	0
Residents	1	3	12	3	0
How well do you think the competency criteria reflect the educational objectives of the program?					
	Very much	Moderately	Little	Not at all	
Faculty	2	5	0	0	0
Residents	5	9	4	1	0

because the process of continuous improvement is inherent to the function of our training program, the current residents and faculty evaluating the process are in fact part of its ongoing examination and modification, and are therefore not objective raters.

Second, responses from graduates, hospitals, and insurers are too few to support any generalizations. Future surveys of these groups should canvas a larger target set. Although the survey was aimed at a broad range of interests, there is no guarantee that the responses reflect that range. Further, concurrent validity represents something of a compromise in attempting to define predictive validity, and our survey only asked the respondents what they thought character-

ized the needs of practitioners. A truer, though impractical, measure would be to examine the actual characteristics of practitioners who are considered experts by consensus.

The results of the evaluation of our program are probably of less importance to psychiatric education than the process of evaluation. Our field has undergone several years of ambitious efforts to define dimensions of competence in psychiatric skills. Now that programs are required to do so in the realms of psychotherapy and general medical competencies, it is valuable to know that methods exist by which we can evaluate the performance and validity of these assessment programs.

**TABLE 3. Responses of graduates and credentialing bodies**

How well do these requirements reflect the skills necessary for the current practice of psychiatry in your setting?				
Extremely well	Very well	Moderately well	Fairly poorly	Very poorly
4	8	2	1	
How well do you believe these requirements reflect the skills that are likely to be necessary for the practice of psychiatry over the next five to ten years?				
Extremely well	Very well	Moderately well	Fairly poorly	Very poorly
	7	8		

## References

1. Wolf, RM: Evaluation in Education: Foundations of Competency Assessment and Program Review (3<sup>rd</sup> Ed.). New York, Westport CT, London, Prager, 1990
2. Monjan S, Gassner SM: Critical Issues in Competency Based Education. New York, Oxford, Toronto, Sydney, Frankfurt, Paris, Pergamon Press, 1979, pp. 1-35
3. Norris N. Understanding Educational Evaluation. New York, St. Martin's Press, 1990, pp. 97-109
4. Hibbard KM, van Wagenen L, Lewbel S, Waterbury-Wyatt S, et al.: A Teacher's Guide to Performance-Based Learning and Assessment. Alexandria VA, Association for Supervision and Curriculum Development, 1996, pp. 33-46
5. Saylor JH: TQM Field Manual. New York, McGraw Hill, 1992, pp. 99-105
6. Weaver CN: TQM: A Step-by-Step Guide to Implementation. Washington DC, ASQC Quality Press, 1991, pp. 137-156.
7. Brainard EZ. A Hands-On Guide to School Program Evaluation. Bloomington IN, Phi Delta Kappa Foundation, 1996, pp. 9-55
8. Marshall J: Assessment during postgraduate training. In JS Gonella, M Hojat, JB Erdmann, JJ Veloski (eds), Assessment Measures in Medical School, Residency and Practice. New York, Springer Publishing Co., 1993, pp. 47-54
9. Hojat M, Gonnella JS, Veloski JJ, Erdmann, JB: Is the glass half full or half empty? A reexamination of the associations between assessment measures during medical school and clinical competence after graduation. In JS Gonella, M Hojat, JB Erdmann, JJ Veloski (eds), Assessment Measures in Medical School, Residency and Practice. New York, Springer Publishing Co., 1993, pp. 137-152
10. Bienenfeld D, Klykylo W, Knapp V: Process and product: Development of competency-based measures for psychiatry residency. Acad Psych 2000; 24:68-76
11. Best Hospitals: PSYCHIATRY <http://www.usnews.com/usnews/nycu/health/hosptl/specpsyc.htm>, 2001