Innovations in testing emerging from the competency assessment movement offer counselors new capabilities in helping their clients to understand aspects of themselves and their problems, as well as to establish directions for development and improvement efforts. New types of tests and measures sample actual behavior more closely than testing instruments previously used: The characteristics they examine are closely linked to performance in a wide variety of jobs, and therefore provide increased focus of assessment on life outcomes. With this new degree of specificity and criterion referencing, implications for counseling, training, and development efforts emerge more clearly than with other forms of testing.

DAVID C. McCLELLAND
RICHARD E. BOYATZIS

Opportunities for Counselors from the Competency Assessment Movement

Until recently, information available from tests has made a limited contribution to the work of the counselor. Such information derives chiefly from tests of academic ability, vocational interest, or specialized disabilities. It has limited utility because test performance does not easily relate to actual functioning in various jobs, nor does it generally give clues about how people can be helped to improve their performance on the job. The competency assessment movement is opening up new opportunities for the counselor by providing more information on job-related competencies in a more explicit, clearly understandable way.

What is competency testing? How are competencies defined, measured, and related to job success? How is the information obtained from a competency test different from the information obtained, say, from an ability test? Why is this new type of information of greater potential use to the counselor? The easiest way to answer such questions is to outline the step-by-step procedures followed in determining the competencies needed for superior performance in a particular occupation. An excellent example is provided by a study by Boyatzis and Burruss (1977) of the characteristics of successful alcoholism counselors in the U.S. Navy. Using this study as an illustration has the added advantage of suggesting some competencies that generally characterize good counselors. Listed below are the five steps in a typical competency assessment process.

FLOW OF THE COMPETENCY ASSESSMENT PROCESS

First, locate some outstanding and average performers in the position. Boyatzis and Burruss (1977) followed the usual practice of asking a number of people who know the job (e.g., counselors themselves or superior officers) to nominate some star performers. When this procedure is followed, inevitably a few people are nominated over and over again, and others are never nominated. We have yet to find an occupation where there is no agreement as to who the really outstanding performers are. It should be noted that this is an *operant technique* in the sense that we do not provide them with names to rate, but instead we call on them to generate the names of the superior performers. As we shall see, *operant measures,* in which a person must operate in the sense of producing an answer spontaneously, are generally more valid and less subject to various testing response biases than are *respondent measures,* in which a person responds to or judges stimuli presented to him or her.

Second, conduct behavioral-event interviews of a small sample of outstanding and average performers. The sample need not be large, because the interviews are simply a method of generating hypotheses as to how these two types of people function differently on the job. Boyatzis and Burruss interviewed 10 outstanding and 16 average U.S. Navy alcoholism counselors. These interviews begin by asking about job functions, about what counselors see their job as requiring them to do. They may even be supplemented with a job element analysis questionnaire (Primoff, 1973) in which the interviewee is asked to rate how important various characteristics are for his or her job.

The main objective of the interview goes far beyond ascertaining what the person thinks his or her job involves to finding out what actual behaviors—including actions, thoughts, and feelings—the person has shown in connection with distinct "behavioral events." Ideally the observer would be like a fly on the wall, so to speak, watching the interviewees function in their daily work. More than that, it would be nice to know what was going on inside the interviewees' heads as they dealt with problems that came up. But short of that, the behavioral event interview attempts to investigate what actions, thoughts, and feelings the person displayed in connection with some vividly successful or unsuccessful events in the past few months. The interviewees are asked to recall particular occasions when they were successful or did their job unusually well, or when they did not do particularly well. They are then asked to describe in detail what gave rise to the situation, when and where it occurred, with whom, and how they dealt with it from beginning to

TABLE 1

Results of a Job Competency Analysis for Navy Alcoholism Counselors

Job Competencies Hypothesized from Interviews	*Test Instrument*	*Test Variable*
Sense of Efficacy	Nowicki-Strickland Locus of Control	Internal Locus of Control
	Picture Story Exercise	Cognitive Self-Definition[a]
Conceptual ability	Test of Thematic Analysis	Thematic Analysis[a]
	Programmed Case Analysis	Number of Correct Predictions[a]
Empathy	Programmed Case Analysis	Improvement Score[a]
	Picture Story Exercise	Caring Motive Profile[a]
Genuineness	Focusing on Feelings Exercise	Focusing Ability[a]
Desire for Personal Growth	Helping Resources	Willingness to Seek Counsel
		Use of Counsel

[a]Significantly associated with job success, indicated by work performance ratings, of counselees 6 months after treatment.

end. In short, the interview is investigative. The interviewer is trying to get as complete a picture as possible of the way the person goes about doing his or her job. Usually about six such episodes, described in detail, are obtained from each person interviewed.

Third, conceptualize the competencies that differentiate superior from average performers. In the course of the interviews, it becomes clear that the superior performers are reporting different kinds of episodes and different ways of responding to them. The task of the psychologist is one of concept formation: What characteristics are shown by superior performers that are not shown by average performers? These traits are called *competencies* because they are characteristics of actual job performance that differentiate superior from average performers.

In defining these competencies, it is very important to also keep firmly in mind the problem of finding a way to measure them. For example, it would not be particularly useful to characterize the superior counselors as "more dedicated" than the average counselors, because psychologists do not have readily available means of measuring degree of dedication. The competencies must also be generic, in the sense of encompassing a wide variety of specific behaviors, and at the same time measurable or potentially measurable. Therefore, we define a *competency* as a generic body of knowledge, motives, traits, self-images, social roles, and skills that are causally related to superior or effective performance in the job.

Table 1 summarizes the results of the job competency assessment analysis carried out for Navy alcoholism counselors. It is shortened somewhat for the sake of simplicity. Notice, first of all, that what might be called threshold skills are not listed in the table. A *threshold skill* is a minimum skill required for the job. In the present instance, for example, it is obvious that the Navy counselors would have to able to read and write, speak English, be loyal to the United States, et cetera. These skills do not differentiate superior from average performers and therefore we do not refer to them as competencies. In the past, threshold skills have often been treated as competencies, and people were selected as better qualified for a job if they had more of these threshold skills, even though more than a certain amount of these skills did not make them better at the job. For example, consider the position of the human service worker. A threshold skill needed for this job is the ability to read bus schedules so that the worker can readily make visits to different parts of the city. We know of at least one civil service jurisdiction (and there are probably many others) in which a test of the ability to read bus schedules is used in such a way that those who can read them faster and better are selected for the job. Yet there is absolutely no evidence that superior human service workers are those who can figure out a bus schedule faster than their peers. In this instance a threshold skill was illegitimately converted into what we call a competency, a behavioral characteristic that differentiates superior from average performers.

To return to Table 1, from comparing behavioral event descriptions, it was hypothesized that the superior Navy alcoholism counselors had demonstrated (a) a greater sense of efficacy; (b) a greater conceptual ability to think in terms of causal patterns in a patient's behavior; (c) more empathy or sensitivity to verbal and nonverbal cues given by a patient; (d) more genuineness in the sense of being aware of their own feelings in a counseling situation; and (e) a greater desire for personal growth.

The fourth step in the process is to find or develop measures for the competencies that differentiate superior from average performers. One possibility is to develop formal coding systems for each of these competencies and use these coding systems to score interviews with new job occupants for any of the competencies. These coding systems must be objective enough so that two judges reading the same interview can come up with the same scores. They can be used to cross-validate the initial competencies identified by blindly coding a new set of interviews with superior and average performers. In the case of the U.S. Navy alcoholism counselors, codes were not developed because the test of the hypotheses about competencies shown was carried out through choosing test instruments designed to measure each of the competencies. They are listed on the right-hand side of Table 1.

The sense of efficacy was measured through a well-known internal locus of control scale (Nowicki & Strickland, 1973). It was also measured through a code that was developed by Stewart and Winter (1974) to score spontaneous stories written to pictures (a version of the classic Thematic Apperception Test) for cogni-

David C. McClelland is a professor of psychology at Harvard University, Cambridge, Massachusetts. Richard E. Boyatzis is President of McBer and Company, located in Boston, Massachusetts.

tive self-definition, or the tendency to think of persons in the story as actively influencing what happens rather than being victims of events. Conceptual thinking was measured by a test of thematic analysis in which the test-takers are presented with two different groups of data. They are asked to determine and describe the differences and similarities between the two groups of data as best they can. The data employed in this study were two sets of three counseling situations derived from the interviews. The quality of the conceptualizations arrived at to show similarities and differences among the situations is scored according to a standard and highly reliable system (Winter & McClelland, 1978).

Empathy was measured in two ways. The Programmed Case (Dailey, 1977) is an indication of the extent to which a person can begin to understand another person as more and more information about that other person is revealed. It consists of a series of 21 episodes from a true life history. Each episode is accompanied by a set of four possible responses that could have been made, only one of which actually was made. Test-takers select the response they think the person made. A specially designed answer sheet indicates whether they are right or wrong, and they are instructed to keep guessing until they find the correct alternative. They are then presented with a new episode and asked what the person did in that situation. Because respondents have more information as a result of each guess, they should improve in their ability to predict what the person will do as they proceed through the case, if they really begin to understand or have empathy for the individual. The number of correct predictions on this test is used as a measure of the respondent's ability to conceptualize the psychological characteristics of the case. Similarly, the degree of improvement from the beginning to the end of the case is considered a rough measure of the ability to empathize with the individual.

Another aspect of empathy is the extent to which a person "feels for" another. To measure this kind of empathy, we used a particular type of motive profile (drawn from the Thematic Apperception Test), which consists of a high score on the Need for Affiliation (see Shipley & Veroff, 1952) and a high score for control or inhibition of activity (see McClelland, 1975). This is called the Caring Motive Profile, because people with these motives generally like being with other people, but in a controlled rather than impulsive way.

Genuineness was measured by a scale for focusing ability, which was a modified version of an instrument designed by Gendlin, Beebe, Cassens, Klein, and Oberlander (1968). Prerecorded instructions ask respondents to explore their feelings about a problem over a 10-minute period. At the end of the 10 minutes, they are asked to answer into a tape recorder specific questions about the experience. Their answers are scored corresponding to the degree they have been able to focus on their own feelings during the 10-minute period. This was considered to be a measure of genuineness, on the assumption that people cannot be honest or real in their behavior unless they are able to readily *experience* and describe their own feelings.

Finally, the desire for personal growth was measured by two questions in which respondents were asked how likely they would be to consult each of 11 different sources of help in case of need and how much they had actually made use of any of these sources of help within the past month, six months, or year.

The fifth step in the process is to administer the tests supposedly measuring the competencies to a new group of individuals rated for job success to see if the tests differentiate the more from the less successful. In the case of the U.S. Navy alcoholism counselors, it was easier than usual to get a measure of job success because the U.S. Navy keeps track of whether those who have an alcoholism problem are able to return to duty or must be discharged. Further, the U.S. Navy monitors the quality of their work once on duty. The score for each counselor was the average work performance ratings of his counselees six months after the counseling had terminated. Twenty-nine counselors whose success ratings were available in this form took the battery of tests described in Table 1. The superior counselors scored significantly higher than the less effective counselors on six instruments measuring 4 of the 5 competencies. Although 87% of the counselees of the 15 superior counselors returned to duty and received satisfactory work performance ratings, only 70% of the counselees of the 14 average counselors had the same success rates. Furthermore, of the superior counselors, 11 had 4 or more of the competencies in high degree, compared with only 2 of the average counselors. In other words, *the competencies hypothesized on the basis of the behavioral event interviews proved for the most part to differentiate the superior from the average counselors as expected,* using most of the particular measures chosen to represent those competencies.

IMPLICATIONS OF THE COMPETENCY ASSESSMENT PROCESS

There were some exceptions, in which tests did not differentiate performers. The measure of locus of control did not differentiate superior from average counselors. In this respect the result is similar to results obtained on other occasions. The locus of control is a respondent measure and therefore is much more likely to be influenced by the person's self-image or by social desirability set than is an operant measure such as cognitive self-definition. In the latter instance, the person is simply writing a story and is less able to control his or her response in a way that will make a good impression. Four of the six measures used in this study that proved valid involved operant responses, in which the counselor had to generate the response that was scored. The other measure that did not prove valid involved self-reports of willingness to seek help to develop counseling skills.

An important point that emerges from studying Table 1 is that the measures are, in many cases, very close to the competencies themselves. The thematic analysis measure involves precisely the kind of psychological conceptualization of behavioral events that a counselor ought to be able to carry out if he or she is functioning well. Similarly, the Programmed Case calls for abilities to understand an individual and to foresee what a person is likely to do based on past behaviors. And focusing ability is a direct measure of the degree to which a counselor is in touch with his or her own feelings.

Thus competency testing comes very close to sampling the actual behavior required on the job. In this respect, it differs markedly from most traditional testing. For example, consider an ability test item such as making a correct analogy. It has always been difficult to understand how being able to complete an item like "Bird is to tree as duck is to ________" is somehow related to performance as, say, a bank teller, policeman, or counselor. Vocational interest tests have the same drawback: As counselors have long known, it is very hard to see how answers to specific items in the test are related to the vocational preferences they indicate. That is, test item answers do not have "face validity," and they have little utility in counseling. How does one explain to a counselee that "disliking people with false teeth" means that he or she is likely to be happy as, for example, a musician or real estate salesman? For most traditional tests, going over actual test responses is of little help in the counseling situation. The opposite is true with competency testing because very often the test behaviors represent the competencies themselves.

A further implication of the fact that test behaviors represent competencies is that competency assessment measures often provide specific directions and guidance for improvement through counseling or training. In other words, the situation is more like taking a driver's test because the test involves using the actual skills that the driver must have to drive well. Thus there is a collaborative relationship between the tester and the testee because failure on the test identifies corrective actions to be taken. Thus, for someone training to be a counselor, it may make a lot of sense to keep performing Programmed Case analyses until that person becomes really good at understanding people and predicting what they will do. It may make good sense to practice the thematic analysis exercise on different materials until one can clearly conceptualize similarities and differences among groups of people.

Because the competency measure, in assessing the behavior itself, usually provides a clear description of the desirable behavior, it can also be used to develop training objectives. Once such objectives are conceptualized, specialized training programs can be developed and evaluated with greater ease than previously as a result of the specificity and clarity regarding the desired behavior.

Furthermore, there is considerable evidence that many of the competencies that have been identified can be directly influenced by education. For example, Winter, McClelland, and Stewart (1980) have recently completed a major research study of how liberal arts colleges are effective in developing a number of competencies, such as the one involved in thematic analysis. Here again we note a difference from traditional testing, in which the goal has been to identify stable abilities or preferences, which are relatively uninfluenced by education or other environmental events.

The job competency analysis for alcoholism counselors in the U.S. Navy is just one example. Many such analyses have been made for various occupations, and a whole new generation of measures of these competencies is in the process of being developed and tested. Table 2 has been prepared to suggest what a few of these other occupations and measures might be. McClelland and Klemp (1974) identified competencies needed for human service work. They designed a special "scenarios test," which consists of a number of taped interviews of human service workers with typical clients. Test-takers listen to the interviews, observe some pictures of the interviewees, and attempt to report correctly what happened in the interview and what advice they would give on a variety of occasions. They are also asked to rate the characteristics of the interviewee; a special "positive bias" subscale of these adjectives clearly differentiated outstanding from average human service workers. Furthermore, a version of the Profile of Nonverbal Sensitivity, developed by Rosenthal, Hall, Dimatteo, Rogers, and Archer (1979), also differentiated superior from average human service workers. Superior human

TABLE 2

Some Representative Competency Measures Associated With Job Success

Occupation	*Competency Identified From Interviews*	*Associated Measure (validated against job success)*
Human service worker	Faith in client's ability to change	Scenarios Test: Positive bias subscale
	Ability to observe and diagnose human problems	Scenarios Test: Casework subscale
		Sensitivity to Nonverbal Cues[a]
U.S. Navy Commanding and Executive Officers	Use of socialized power	Picture Story Exercise: Leadership Motive Profile[b]
	Plan, formulate policy	Skill-in-Scheduling Test
	Optimize results in conflict situation	Managerial Style Test: Optimizes
Organizational consultant	Concept formation ability	Learning Style Inventory and Test of Thematic Analysis
	Empathy	Picture Story Exercise: Caring Motive Profile
Foreign Service Information Officer	Political judgment	Critical Decisions Test
	Empathic sensitivity	Sensitivity to Nonverbal Cues[a]

[a]Correctly diagnosing emotions presented in content-filtered, taped speech segments, from Rosenthal et al.'s Profile of Nonverbal Sensitivity (1979).
[b]Also shown to significantly differentiate superior from average managers in industrial organizations and other public sector organizations ($N = 450$) (Boyatzis, in preparation).

service workers were better able to diagnose what emotions were being expressed in a short taped speech fragment, even though the words could not be understood.

To shift to a different example, among U.S. Navy commanding and executive officers as well as managers in large corporations, it has repeatedly been found that the Leadership Motive Profile is more characteristic of the more successful managers. This profile shows a high Need for Power (Winter, 1973), higher than the Need for Affiliation, and a high index of self-control or activity inhibition (Boyatzis, in preparation). In other words, successful managers must be interested in influencing others, not too easily sidetracted by their liking for others, and have good control over themselves. Furthermore, superior U.S. Navy commanding and executive officers show greater skills in scheduling and in managing in ways that optimize the outcome in a conflict situation.

In a competency study of organizational consultants, McClelland (1975) reported that the more effective consultants showed a greater ability to detect the problems in a client organization than did their less effective colleagues who examined the same information. This ability was assessed through the Learning Style Inventory (Kolb, 1976) and the Test of Thematic Analysis (Winter & McClelland, 1978). The superior consultants also showed a greater concern for their clients as measured through the Caring Motive Profile.

As a final example, McClelland and Dailey (1973) reported that superior U.S. Foreign Service Information Officers show greater political judgment as indicated by making the right responses to a series of critical episodes drawn from behavioral event interviews. They also show greater sensitivity to paralinguistic cues in the auditory research of the Profile of Nonverbal Sensitivity, like the superior human service workers mentioned above.

THE VALUE TO COUNSELORS OF THE COMPETENCY ASSESSMENT PROCESS

The examples discussed in this article have been taken from a much longer list, but they should be sufficient to indicate that the competency assessment movement is in fact spreading a fairly wide net. It is covering more and more occupations and developing a longer and longer list of competencies that can be assessed through either interviews or tests. These competency measures should be particularly valuable to counselors for a variety of reasons. In the first place, they are more clearly job-related than are most traditional tests. They should be of particular value in vocational counseling because many of them involve generic skills, such as the empathy measures, that are required for a number of different occupations.

In the second place, the new competency measures provide a much wider variety of test types and responses than do the traditional paper-and-pencil multiple-choice tests. Thus the counselor has a much longer list of valuable types of information about how the counselee is behaving in a variety of situations.

In the third place, the new competency measures include many operant measures. These measures seem to be more valid in making critical distinctions regarding performance in jobs in a nondiscriminatory manner than many other types of tests.

In the fourth place, the counselor is in a better position to feed back useful information to the counselee than with traditional types of test scores. The information is directly helpful in defining what new responses the counselee should learn if he or she is to gain the competence in question. For example, it is now known that there are particular psychological education courses that increase a variety of competencies, such as self-definition (deCharms, 1976), achievement motivation (McClelland, 1972), nonverbal sensitivity (Rosenthal et al., 1979), thematic analysis skills (Winter, McClelland, & Stewart, 1980), and so on. As we become more knowledgeable as to just what competencies are needed for various positions, it will be easier to design special courses aimed at teaching them.

It is our feeling, in short, that the competency assessment movement will greatly increase the variety and precision of the psychological assessment devices available to counselors, who will then be able to do their job much better and play an increasingly important role in the educational process.

REFERENCES

BOYATZIS, R. E. *The competent manager: Exploration into a generic model of management.* Boston: McBer & Co.. in preparation.

BOYATZIS, R. E., & BURRUSS, J. A. *Validation of a competency model for alcoholism counselors in the U.S. Navy.* A final report submitted to the U.S. Naval Alcohol Rehabilitation Center, San Diego, California. Boston: McBer & Co., 1977.

DAILEY, C. *Assessment of lives.* San Francisco: Jossey-Bass, 1977.

deCHARMS, R. *Enhancing motivation: Change in the classroom.* New York: Irvington-Halsted-Wiley, 1976.

GENDLIN, E. T.; BEEBE, J., III; CASSENS, S.; KLEIN, M.; & OBERLANDER, M. Focusing ability in psychotherapy, personality, and creativity. *Research in Psychotherapy,* 1968, *3,* 217-241.

KOLB, D. A. *Learning style inventory (technical manual).* Boston: McBer & Co., 1976.

MCCLELLAND, D. C. What is the effect of achievement motivation training in the schools? *Teachers College Record,* 1972, *74,* 129-145.

MCCLELLAND, D. C. *A competency model for human resource management specialists to be used in the delivery of the human resource management cycle.* Report prepared for the U.S. Navy Human Resource Management Center. Boston: McBer & Co., 1975.

MCCLELLAND, D. C., & DAILEY, C. *Evaluating new methods of measuring the qualities needed in superior Foreign Service Information Officers.* Report prepared for the U.S. Information Agency. Boston: McBer & Co., 1973.

MCCLELLAND, D. C., & KLEMP, G. O., JR. *Validation of a human service worker test.* Report prepared for the Massachusetts Civil Service. Boston: McBer & Co., 1974.

NOWICKI, S., & STRICKLAND, B. R. A locus of control scale for children. *Journal of Consulting and Clinical Psychology,* 1973, *40,* 148-154.

PRIMOFF, E. *How to prepare and conduct job element examinations.* Washington, D.C.: U.S. Civil Service Commission, 1973.

ROSENTHAL, R.; HALL, J. A.; DIMATTEO, M. R.; ROGERS, P.; & ARCHER, D. *Sensitivity to nonverbal communications: The PONS test.* Baltimore: The Johns Hopkins University Press, 1979.

SHIPLEY, T. E., JR., & VEROFF, J. A projective measure of need for affiliation. *Journal of Experimental Psychology,* 1952, *43,* 349-356.

STEWART, A. I., & WINTER, D. G. Self-definition and social definition in women: *Journal of Personality,* 1974, *42,* 238-259.

WINTER, D. G. *The power motive.* New York: The Free Press, 1973.

WINTER, D. G., & MCCLELLAND, D. C. Thematic analysis: An empirically derived measure of the effects of liberal arts education. *Journal of Educational Psychology,* 1978, *70,* 8-16.

WINTER, D. G.; MCCLELLAND, D. C.; & STEWART, A. J. *Competence in college: Evaluating the liberal university.* San Francisco, Calif.: Jossey-Bass, 1980.