In Berger, L. A. & D. R Berger Eds (2003) *The Talent Management Handbook*, New York NY: McGraw-Hill

Chapter 6.

HOW COMPETENCIES CREATE ECONOMIC VALUE

Lyle M. Spencer PhD President

Spencer Research & Technology

This chapter describes empirical methods for developing competency models that meet professional and legal reliability and validity criteria. Specific steps and templates are provided for quantitative calculations.. A case study of competency model development, business case, application and evaluation is provided.

Definition

A competency is a *reliably measurable, relatively enduring* characteristic (or combination of characteristics) of a person, team or organization, which causes and statistically predicts a *criterion level of performance.*

Italicized words and phrases in the definition have technical meanings which have important implications for building competency models

Reliably measurable

Reliably measurable means two or more independent observers or methods (tests, surveys) agree statistically (usually r =.80) that a person demonstrates a competency characteristic. Inter-rater reliability is critical to insure the competency characteristic is a consensual construct, measure accurately, and avoid bias.

Professional standards, scientific practitioner ethics and the law require reliable and valid measures

- Professional Standards

The Standards for Educational and Psychological Testing, prepared by a committee of the American Educational Research Association, American and Psychological Association, and National Council on Measurement in Education (1999), require measures (and by inference, human resource programs based on these measures) to be reliable and valid (that is, to statistically predict) outcomes of (economic) value to individuals or organizations.

Competency researchers and HR practitioners are regularly savaged by critics for failing to publish reliability and validity data: for example, Barrett (2000) denounces competencies as "slickly packaged junk science perpetrated by unscrupulous consultants on ignorant customers." Published data about the efficacy of competency programs exist (see Chapter Nine), but competency advocates have largely failed to bring these data to human resource (HR) professionals' attention.

- Legal Requirements

U.S. and Canadian courts, under civil rights and (in Canada) pay equity laws, have ruled that "any [HR] decision-making processes, from background checks to supervisory performance ratings, that affect an employee's status in an organization, are tests, and thus subject to scrutiny for adverse impact" (Latham & Wexley, 1981). These rulings effectively extend requirements for statistical reliability and validity to any assessment for selection or promotion, any development opportunity and any performance appraisal affecting pay or career opportunities.

Legal requirements for scientific reliability have been expanded by U.S. Supreme Court Associate, Justice Stephen Breyer's decision for the majority in *Kumho Tire, Inc.* v. *Carrnithael* (119 Sup. Ct. 1167 [1999]), which extends an earlier U.S. Supreme Court ruling in *Daubed* v. *Merrell Dow Pharmaceuticals, Inc.* (509 U.S. 579 11993]). Daubert required expert witness testimony to be based on "tested scientific knowledge, demonstrate reasonable reliability criteria, have been subjected to peer review, report the size of the known error rate for findings . . [and] establish whether the knowledge enjoys widespread acceptance in the scientific community" (Daubert, cited in Wiener, 1999).

Valid development opportunities, for example, can clearly make a difference in an employee's status, and for this reason they have been the subject of many legal battles (such as the 1978 *Bakke* v. *Regents of the University of (California)*. Access to (quality) COMPETENCY education and training opportunities almost certainly falls under these laws. An employee can complain: "You sent me to the

'feel-good' course when my colleagues got to go to *validated* training which helped them show improved business results and get promoted? Discrimination!" And lawsuit?

The legal status of psychological tests and programs in European Community countries under EC and individual country labor laws and union and worker council agreements is less clear, but many observers believe scientific validity requirements for HR practices will become law in Europe. Multinational HRIS vendors (for example, PeopleSoft and SAP) are designing their systems to provide data on whether competency programs pass legal tests of reliability and validity.

- Economic Value Added

Evaluation methods that look at the economics of human resource programs are premised on the same survival-of-the-fittest concept that governs all businesses: that is, the goal is to help investments flow from less valuable uses to uses where they generate the highest returns.

Economic value-added (E\A), cost-benefit, and return on investment (ROI) analyses lead HR staff to improve practices by helping them to

- Focus on the *right* problems or opportunities—those with the greatest cost or value, respectively to the firm.
- Focus on *interventions* that will have the maximum impact on costly problems and valuable opportunities.

Demonstrating the economic value of outcomes also enhances the professional longevity credibility, and satisfaction of COMPETENCY researchers and practitioners in several ways.

First, the HR function competes with every other organizational function for capital investment funds. HR professionals are more likely to he able to convince their customers to adopt programs when they can describe program benefits in economic terms. Investment proposals with business cases showing compelling ROI projections are more likely to be funded. “Soft" programs and staff (that is, those lacking economic justification are more likely to be cut. Second, HR programs are increasingly emphasized in making ISO 9000, JACHO, Deming, arid Baldrige audits and awards. Most of these assessments arc qualitative. Economic value—added data can provide powerful measures of HR programs’ quality Hard data showing that HR interventions made a meaningful business contribution to an organization are more likely than oilier evaluations to find their way into management reports and personnel folders and to enhance HR staff careers.

Finally, all advanced methods in human resources: value chain analyses, development of business cases based on the probable shift in competency-based performance, he was in enterprise resource programs (SAP, PeopleSoft HR technologies, human resource asset accounting, depend on validated statistical relationships between competency predictor variables and economic outcome variables accounting.

Relatively enduring

Relatively enduring means a competency measured at one point in time is statistically likely to be demonstrated at a later point in time. For example, a candidate selected for competencies "achievement motivation" or "initiative" now will demonstrate these behaviors three (or 30) years from now. Stability of competency traits is critical for predictive validity—predicting how an employee will behave in future jobs.

Competency *characteristics*

Competency *characteristics* are content knowledge, behavior skills, cognitive processing (IQ), personality traits, values, motives, and occasionally other perceptual or sensori-motor capabilities (reaction time for combat pilots, taste and smell for sommeliers), which validly predict performance outcome criteria. Several hundred competencies have been identified, but 20-25 account for 85% of the variance in most jobs

Knowledge, skill and personality characteristics required for minimally acceptable performance are called "*threshold*" competencies—those which distinguish people who can do the job from those who cannot. Characteristics that predict superior performance are "*differentiating*" competencies because they statistically differentiate superior from average performers.

Validity

"A difference which *makes* no difference *is* no difference."

William James, 1842-1912

Harvard professor,, "Father of American Psychology"

Validity means that a reliably measured competency predictor (independent) variable statistically predicts a criterion performance (dependent) variable.

statistically predicts

Predictor (independent) variables ========➔Criterion (dependent) variables

Competency 1 — Performance: $ or other quantitative

Competency 2… — results variable(s)

etc.

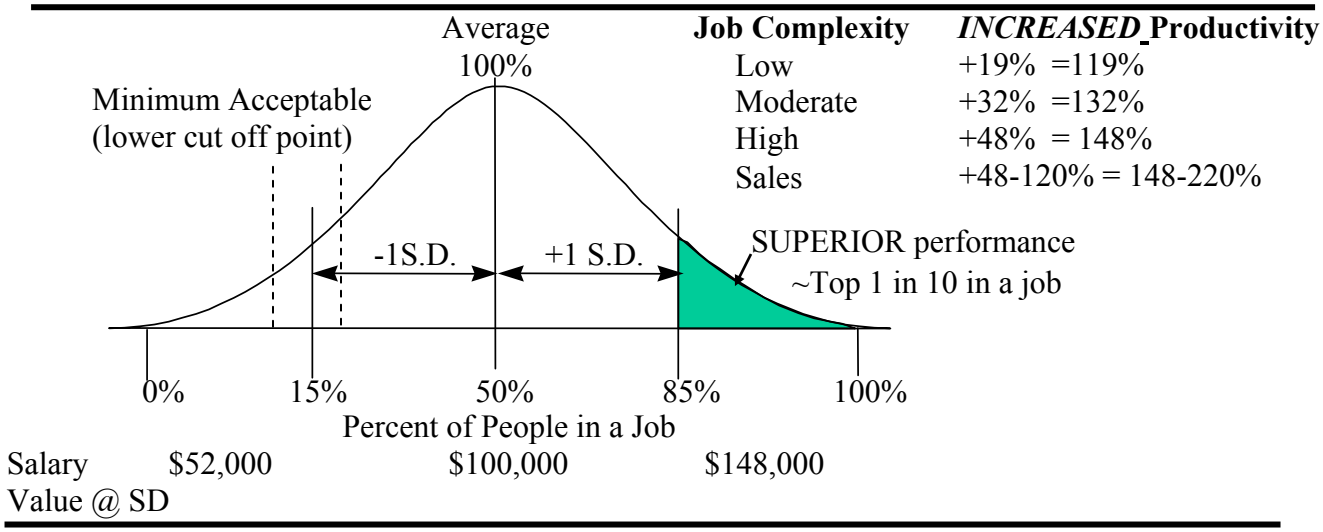
It cannot be to strongly emphasized that "competence" or "talent" as a concept has meaning *only* to the extent it empirically predicts a criterion level of performance. Very few competency models developed in most organizations meet this standard. Competency lists or models lacking empirical validation are meaningless and useless. To paraphrase William James "A 'competency' which doesn't predict measurable performance *is no* competency."

Criterion level of performance

Criterion level of performance is a point on a normal distribution (bell curve—see Figure 1) of the performance results of all employees in a job. This point can be anywhere on the curve, but the three most common reference points: are

- *minimally acceptable*: the lower cut point, below which employees will be replaced because their performance is below what the organization can tolerate
- *average* performance: the mean of the bell curve
- *Superior* performance, one standard deviation above the mean (top 15%, or roughly top 1 of 10 employees in a job.

Figure 1. Economic Value Added by Superior (+1 SD) Performance



The superior performance point on the curve is the most useful for improving performance, for 3 reasons

- it has *known economic value* added, which increase is with the complexity of job. meta-analytic studies have shown that in low complexity (semiskilled labor and clerical) jobs, superior performers plus one standard deviation above the mean are 19 % more productive than averages. In moderate complexity jobs (technical and

first level supervisory) superiors produce 32% more than averages. In complex (professional, managerial and executive) jobs, superiors are 48% more productive than averages, i.e. produce as much as 1.5 average employees. In sales positions superior performers sell from 48 -- 120% more than averages; the usual rule from his the best sell twice what the averages do.

- it defines the *most widely used metric, "effect size", for measuring the impact of HR interventions.* Human resource programs selection, training, performance management, goal setting, feedback, coaching add value by shifting employees from wherever they are performing on the bell curve toward superior performance. Fractions and multiples of the one standard deviation difference between average and superior performers is a common yardstick for how much difference a human resource program: can make, for example "training T increased performance .40 effect size."
- it is a *benchmark* which drives individuals and HR to improve performance to point *better* current average. Individuals or HR departments NOT using a performance benchmark better than their current (average) level of achievement select, training, performance manage to their current level of *mediocrity.*

Competency modeling research is akin to benchmarking best practices. Benchmarking has two stages: benchmark *"what":* measures of superior performance: lowest cost per unit, fastest cycle time, highest quality, largest profit margin of best performing firms. Benchmark *"how"* follows: the study of the processes, materials, equipment, and human resources the best performing organizations use to get a superior results.

The power of benchmarking depends on its "benchmark what" criterion sample: the economic value added by best results from best practices. Nobody benchmarks mediocre performance, except to use as a contrast to highlight what best practices make the difference.

Benchmark *"how"* in competency research is the study of superior performers to identify characteristics and behaviors they use to achieve the "what" criterion that averages do not. This approach is called "discovery of grounded theory" using "extreme samples"—which means predictor variables are identified in a small sample of people getting outstanding (extreme) criterion results—in this case, the best, most superior performers. The Predictor variable ➔Criterion variable direction is reversed: study of people in the Criterion performance sample tells researchers what the Predictor competency variables are.

Predictor Variable		Criterion Sample
	inference	
Superior Performers' competencies	<==========	Superior performance

Steps in developing a valid competency model

The Steps in developing a competency model are

1. Define Performance Criteria
2. Analyze Criterion Sample
3. Collect data
4. Develop Competency Model
5. Validate Competency Model

6. Application
7. Evaluation

A seven step protocol for developing competency models is shown below with a recent case study example. "Incon" is a US$5.5 billion industrial controls firm with 400 branch managers (BMs) in North America. In 1997 Incon developed a BM competency model for a pilot group of 98 US BMs, applied the model to develop a BM training program, and validated the model against branch profits 1998-2000.

Step 1: Define Performance Criterion

The initial step of defining the performance criterion appears obvious, but in many organizations, managers lack consensus about measures of output performance. Most firms have explicit or de facto *balanced scorecard* variables, but these need to be probed to determine what management really values.

Case Example

Incon's balanced scorecard for sales managers included growth in revenues, return on sales, cost reduction, customer satisfaction, improvement in productivity and operational efficiency, sales of new products and services, organizational climate, and qualified turnover of subordinate managers and salespeople. Balanced scorecard metrics can be aggregated by multiplying the

weight for each measure by an individual or group's score on it. In Table 1 example below, weights have to equal 100, and scores are –2 to + 2 standard deviations in performance on the variable rated. In the example, weights and scores for all measures produce a weighted total score of 60, out of a possible score of 200 (100 weights x +2 maximum score). A superior criterion sample could be identified by taking the top10 or 15% of managers on the total measure score.

Table 1. Balanced Scorecard Criterion from Multiple Weighted Measures

Measure	**Weight**		**Score -2 to +2**	**Criterion Score**	
Organizational Climate	*5*		0	*0*	
Qualified turnover of managers and salespeople	*10*		-1	*-10*	
Productivity Improvement	*5*		2	*10*	
Sales of new products and services	*10*		1	*10*	
Customer Satisfaction	*10*		1	*10*	
Growth in revenues	*20*		0	*0*	
Profit Margin	*20*		1	*20*	
Growth in profits	*20*	**100**	1	*20*	**100**
TOTAL		**100**		*60*	**100**

nn= mean Balance Scorecard Ratings developed by HR focus groups

nn (bold) = valuation by top management

In reality, when pressed, finance told HR that the only performance measure that mattered was increased profits: growth in revenues x return on sales. This measure was used as the dependent variable in developing the business case and evaluating the competency model program.

Step 2. Analyze Criterion Sample

The first question in developing a business case should always be: is there enough variance in the value of dependent variable to make investment in an intervention worthwhile. This question can be answered by calculating the economic value of the problem or opportunity the HR program will address by

- Valuing the problem: the cost per problem incident x # incidents

  and/or

- Valuing the opportunity: finding the economic value added (EVA) per employee (team or firm) per year at the benchmark or desired level of performance-----for example, a criterion sample of' employees (teams, firms) + 1 SD above the mean (EVA/employee/year x # employees).

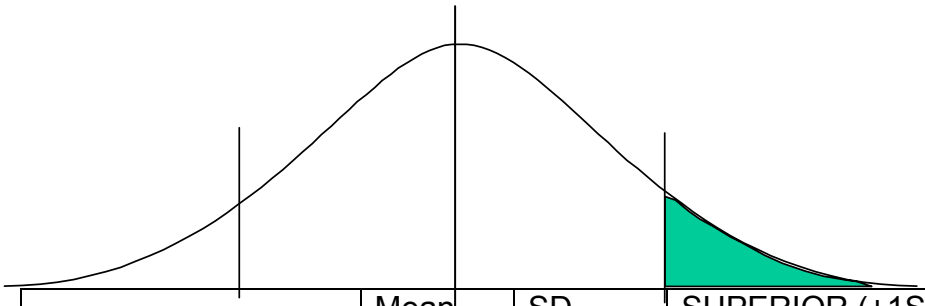
CASE EXAMPLE

Data for the business case were easily developed from Incon financial records with basic descriptive statistics. Sales for branch managers ranged from $4 million to $90 million, with a mean of $17.0 million. As shown in Figure 2., BMs one standard deviation above the mean had 5.66% higher return on $12.8 million more sales, worth $2.94 million in yearly profits, 134 percent more than average

performers. Variance in BM Performance is very large, hence offers a large opportunity for an HR intervention that improves average BM performance.

Figure 2. Branch Manager Variance in Economic Value Added.

CRITERION SAMPLE: AVERAGE V. STAR (+1 SD)
BRANCH MANAGER VARIANCE IN ECONOMIC VALUE ADDED.



	Mean	SD	SUPERIOR (+1SD)
Revenue	$17.02M	$12.82M	$29.84M
Profit Margin	5.60%	5.66%	10.99%
Profit	$1.32M	**$1.69M**	$2.94M

The second business case question is how much impact is the HR intervention likely to have on business results dependent variable(s) i e its probable EVA in problem cost savings or opportunity increased profits. This question is answered by finding (a) the validity of the competency model, i.e. how much of the $1.69M difference BM competencies HR might improve account for, and (b) the power or effect size the HR intervention under consideration have to improve competence.

Step 3. Collect Data

Numerous methods can be used to collect data to build competency models. The challenge is getting reliable and valid data within a reasonable time and at a reasonable cost. As with most investments, more valid methods tend to cost more and

take more time, but returns in increased validity of HR programs justify these investments.

a. Behavioral Event Interview (BEI). The BEI asks interviewees about the most critical situations they have faced on their job: peak high points and peak low points. Probes are very specific but nondirective:

   i. What was the situation--which can be scored by analysts for what situations interviewees considered critical (often very different for superior and average performers).

   ii. What lead up to the situation-- coded for whether the interviewee initiated or reacted to the situation, the extent of which he or she sought information to anticipate or prepare for events.

   iii. Who was involved in the situation--coded for the breadth of people the interviewee typically interacts with:: no one, peers, top management, external experts, customers, in what priority order.

   iv. What the interviewee observed, saw heard, read paid attention to in the situation—coded for information seeking and pattern recognition.

   v. What the interviewee felt and wanted to have happen in responding to the situation--codeable for emotional self-control, self-confidence, and motivation.

   vi. What the interviewee thought about in analyzing or understanding the situation and deciding whether and how to respond to it--

coded for conceptual and analytical thinking and strategic planning.

vii. What the interviewee actually did-- coded for behavioral skills, initiative

viii. what the outcome of the situation was--coded for what a interviewee thinks is an important outcome (achievement, conciliation, achievement, influence).

The BEI provides a wealth of narrative data which can be used to develop case studies and training exercises showing what the best performers actually do in the most critical situations they face. The BEI is open ended so analysts can discover and code new competencies as well as those in existing dictionaries. The BEI consistently shows the highest reliability and validity in predicting the future employee performance. The BEI's drawback is that it requires skilled interviewing and reliable coding, hence is expensive because professional labor intensive.

b. Survey: two survey methods are commonly used

i. Rating of competencies required by a job by managers, and sometimes superior incumbents, on importance and frequency for superior and adequate performance; trouble likely if employees lack the competency, etc.

ii. 360 degree ratings of superior, average and minimally acceptable (or a full normal distribution of) employees in a job by managers, peers, subordinates,

customers and others who have an opportunity to observe employee job performance.

Survey data are relatively quick and inexpensive to collect especially with Internet technologies. Surveys however measure only those competencies that the survey makers already know about, and cannot discover new competencies. Validity of managerial and subordinate ratings of subject competencies are r = .35-40 (r^2 =.10 -- .15), indicated validated 360^o surveys are a valuable measurement tool.

c. Expert Panel/focus group. This approach usually involves a 10 to 20 managers, human resource specialists, academics, sometimes incumbents and others who brainstorm or "guestimate" what they *think* the competency characteristics required for superior minimally acceptable performance are. Twenty years of research have shown that at best, focus groups are 50 percent accurate-- but competencies they identify are not defined specifically enough to be reliably coded. As shown in Table 2. 25 percent of what panels identify are "folklore" virtues and values which are not confirmed by validation studies. Another 25 percent of competencies which do validate are missed by focus groups, because panel members lack psychological vocabulary for identifying valid competencies.

Focus groups are useful for involvement organization members and marketing competency research efforts, but are rarely useful (unless accompanied with critical incident data) for identifying valid competencies.

Table 2. Panel v. Validation Study Accuracy in Identifying Competencies

PANEL	VALIDATION STUDY
•***25% Folklore Virtues***	***Not valid***
50%-- ***but too vague to code reliably*** •**Comp1** •**Comp2** •**Comp3**	**50%--*precise: coded reliably for frequency, level*** •**Comp1** •**Comp2** •**Comp3**
Panel Misses 25%	•**Comp 4** •**Comp 5** •**Comp 6**

d. Job Competency Data bases

Databases built into advanced human resource information systems and enterprise resource programs (ERP) include expert system engines which, given the requirements of a job, can look through all similar requirements identified in competency studies to identify competencies which predicted high-performance in jobs with these requirements. For example, an American telecommunications company wanted to expand sales to Eastern Europe. Large telecommunications systems sales in European countries are highly centralized under governmental ministers of telecommunications. The firm had never had any high-level people in Europe, so had no criterion sample of salespeople from which to identify competencies for this job. A successful salesperson in European telecommunications markets requires a combination of diplomatic and senior high-tech sales/account manager competencies. An

HRIS consultant's database data included several competency models for diplomats and high-tech senior level account managers. The database engine was able to infer from a job requirement the competencies likely to predict success in the job. And shown in the Table 3 should

Table 3. Competencies identified by Expert System for High-level, high-tech sales in Eastern Europe

Job	Task Requirement	Competencies needed
Diplomat	Contact, lobby persuade high level foreign government officials	+Cross-cultural empathy Level4 +Speed of learning political networks (organizational awareness) Level 6 +Relationship Building Level 7
High Tech High Level Account Manager	Strategic advice to top management Lead complex systems integration projects	+"Comfortable in Boardroom" Influence Skills Level 4 +Strategic thinking: IT and Communications systems for LT sustainable competitive advantage Level 4 +Customer Service Orientation Level 7 +Team Leadership in complex systems integration Level 7

e. Observation of on-job behaviors trained observers watch superior and average employees, listen in on phone calls, etc. and code observed behaviors for competencies. Observation is a valid research method but rarely used due to expense, and the fact that equally valid data can be obtained from BEI transcripts.

3. Other assessment methods such as assessment centers, psychometric tests work samples can be used for identifying competencies and by comparing the results of superior performers to average performers or minimally acceptable performance to those incapable of doing a. job. These methods again find only what test items or simulations measure. Assessment centers equal BEIs in having high validity (r= .5-.6, r2 = .25-35) but are more expensive (1 PhD coder for every two assesses) so are rarely used for competency model development v. a competency-based selection or developing application. Psychometric tests are generally used for screening for lower-level jobs (intelligence, vocabulary, language, typing. other sensori-motor skills.)

Step 3. Building Competency Models

The process of developing competency models is essentially

a. identifying those characteristics which competencies which statistically distinguish superior from average performers; then
b. arranging valid competencies in a model that is most easily understood by manager, human resource professional, and employee users.
c. Identifying competencies which predict significantly greater economic value added

Data from surveys are easily scanned in and summed for use in basic statistical tests. Interview data must first be coded competency by level and frequency for superior and average performers. A 1.5 hour BEI (80 page transcript) contains about 80 codable phrases. Coders use a competency dictionary (see Spencer and Spencer, 1993).

Initial competency models were "one offs:" each model and the competencies which described it were unique. Researchers quickly realized, however, that some "core" competencies e.g. achievement motivation or accurate empathy appeared again and again in models for many job families. But because no standard "language" or dictionary for competence existed, these competencies were called many different names by researchers in different organizations and countries. For example, achievement motivation was known by 20 different labels: "Concern for Results," "Performance Orientation," et.al. To create competency dictionaries from meta-analyses of many competency models, a "lowest common denominator" language which could capture in similar terms data in all competency models and translate any competency definition into any other. This realization led to the development of an "atomic" approach to competency classification.

- **Competency "atoms" I**n philosophy, atoms are the smallest, indivisible units of matter or ideas, e.g. protons, neutrons and elections (currently quarks) in physics. In psychology an atom is perhaps best defined as "the smallest useful unit of observation." This unit differs by the type of psychological

analysis. For example, for competency research, an atom is a *behavioral indicator* e.g. "wants to *do better,*" an atom in the "element" achievement motivation.

A second measurement issue is the "**Quanta"** strength or "energy" of competency variables. "Quanta" in physics are discrete, whole number units of energy (orbital levels of electrons; higher orbits equal higher energy.) Many behavioral science methods have been developed to measure variable strength. Examples are the sum of item scores (e.g. Likert 1= low to 5 = high), items correctly answered e.g. vocabulary words or analogies on an intelligence test, frequency (the number of times a competency is coded in an interview transcript), and Just-Noticeable-Difference Scales

An important finding of the research used to develop competency dictionaries was significant variation in the weight or strength of examples of the same competency drawn from different models. Some "Achievement" stories seemed much stronger than others, some examples of Analytical Thinking were much more complex than others.

When verbatim examples of each competency were collected from a variety of jobs and Q-sorted by a number of judges according to the extent to which they indicated more or less of the competency in question, competency examples were found to have scaling properties: a clear progression from lower to higher levels on four dimensions:

- *Intensity* of the intention (or personal characteristic) involved or *completeness of actions* taken to carry out an intention.
- *Complexity*: taking more things, people, data, concepts or causes into account..
- *Time horizon*: seeing further into the future, and planning or taking action based on anticipation of future situations; e.g., acting now to head off problems or create future opportunities.
- *Breadth of impact*: number and position of people impacted, e.g., on a scale from a subordinate or a peer to the CEO of the organization, to national or international leaders; or the size of the problem addressed, e.g., from something affecting part of one person's performance to something affecting the entire organization.

For example, Table 4 shows a 6 level scoring for Achievement Orientation. Levels are descriptions of behavior "just noticeably different" in increasing intensity or completeness of action, complexity, breadth of impact, and time horizon. An example of coding a phrase in a interviewee transcript is:

"Whom when I took over, efficiency was about 85 percent; now its up to 96 percent-- and it's saved us several million dollars a year" (Achievement level 4: cites before v. after improvement with quantitative figures).

In Berger, L. A. & D. R Berger Eds (2003) *The Talent Management Handbook*, New York NY: McGraw-Hill

Table 4. Achievement Orientation (ACH)

Just-Noticeable Difference (JND) Scale

Achievement Orientation: A concern for working well or for surpassing a standard of excellence. The standard may be one's own past performance (striving for improvement); an objective measure (results orientation); outperforming others (competitiveness); challenging goals one has set; or even what anyone has ever done (innovation). Unique accomplishment also indicates ACH.

1. **Wants to Do Job Well**: Tries to do the job well or right, or meet objectives set by others e.g. make a sales quota or budget. May express frustration at waste or inefficiency (e.g., gripes about wasted time and wants to do better) but does not initiate specific improvements.

2. **Creates Own Measures of Excellence**: Develops specific methods of measuring outcomes against a standard of excellence not imposed by others. May focus on new or more precise ways of meeting goals set by management.

3. **Improves Performance**: Makes specific changes in the system or in own work methods to improve performance (e.g., does something better, faster, at lower cost, more efficiently; improves quality, customer satisfaction, morale, revenues), without setting any specific goal.

4. **Sets and Works to Meet Challenging Goals**: "Challenging" means there is about a 50-50 chance of actually achieving the goal—it is a definite stretch, but not unrealistic or impossible. OR refers to specific measures of baseline performance compared with better performance at a later point in time: e.g., "When I took over, efficiency was 85%—now it is up to 96%."

5. **Makes Cost-Benefit Analyses**: Makes decisions, sets priorities or chooses goals on the basis of calculated inputs and outputs: makes explicit considerations of potential profit, Return-on-Investment or cost-benefit analysis. Analyzes for business outcomes. (To code, the person must show: 1)specific mention of costs and 2) specific benefits and 3) a decision based on the balance between them.

6. **Takes Calculated Entrepreneurial Risks**: Commits significant resources and/or time (in the face of uncertainty) to increase benefits, (i.e., improve performance, reach a challenging goal, etc.

Often a combination of frequency and level is the best screen for competencies and levels likely to differentiate superior from average performers. Table 5 shows achievement motivation frequency times level

scorers for partners, partner selectees and non-partners in a worldwide accountancy firm. In this case the screen used to identify competency levels was t-test probability <.05, and r^2 > .05 (5 percent or more of the variance in manager performance)-- cells highlighted in Table 5. Achievement levels 1 and 2 are infrequently scored and did not differentiate. Levels three, four and five statistically significantly discriminate on the t-test or account for more than 5 percent of the variance between partners and non- partners. These three levels are candidates for entry in a multiple analysis of variance to develop a predictive equation for selecting partners. Note levels six of achievement is not shown by any of the managers. This pattern is typical using just those with different skilled competency coding systems: lower levels don't differentiate and higher levels are scored too infrequently to be statistically reliable.

Table 5. Achievement Frequency x Levels with Differentiate Partners and Partner Selectees from Non- Partners (N=16)

Achievement Level	1	2	3	4	5	6
Ptr freq	0.0	0.3	1.5	0.3	1.7	0.0
Ptr freq*lvl	0.0	0.7	4.5	1.3	8.3	0.0
PtrSlt freq	1.0	0.0	2.0	0.5	3.5	0.0
PtrSlt freq*lvl	1.0	0.0	6.0	2.0	17.5	0.0
NP freq	0.0	0.0	0.8	0.6	1.2	0.0
NP freq*lvl	0.0	0.0	2.4	2.4	6.0	0.0
dif	1.0	0.7	8.1	0.9	19.8	0.0
t-test p<	0.33	0.33	0.02	0.13	0.04	#DIV/0!
r2	0.01	0.03	0.22	0.10	0.08	#DIV/0!
r	-0.08	0.18	0.47	0.33	0.28	#DIV/0!

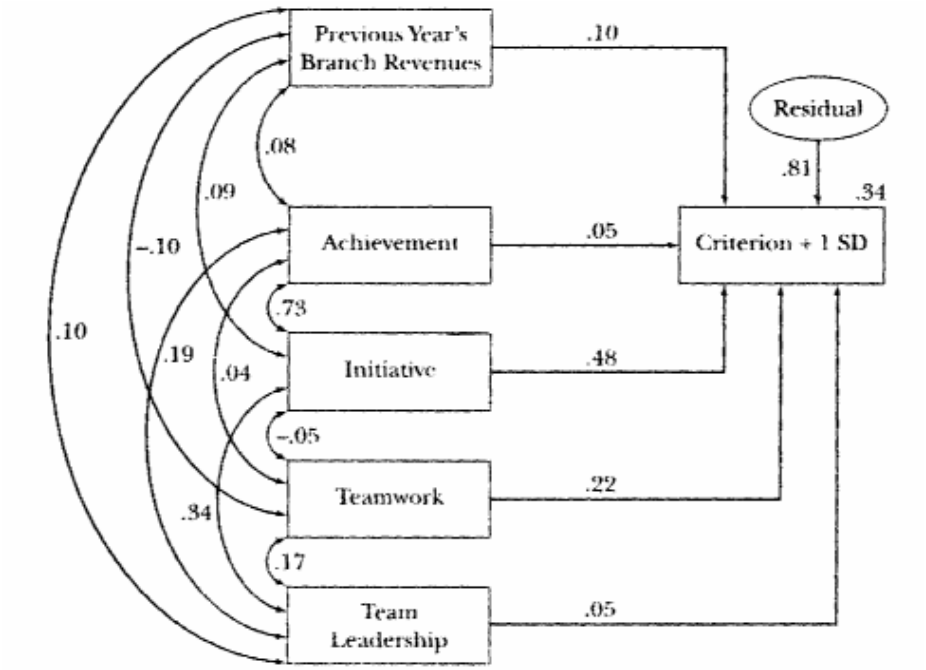
Levels of Achievement competency which statistically significantly differentiate Partners and Partner Selectees from Non-Partners (t-test p<.05 and/or r2 >.05)

In Berger, L. A. & D. R Berger Eds (2003) *The Talent Management Handbook*, New York NY: McGraw-Hill

CASE EXAMPLE

Incon completed a competency study of Branch Managers and used a sophisticated regression method called structural equation modeling to identify competencies and competency interactions which predicted economic performance. The model accounted for 34% (r2) of the $1.69M +1 standard deviation effect size difference between superior and average BMs. As shown in Figure 3, competencies that differentiated superior performers included Achievement (ACH), Initiative (INT), Teamwork (TW) and Team Leadership (TL).

FIGURE 3. COMPETENCIES PREDICTING +1SD SUPERIOR ECONOMIC PERFORMANCE AMONG U.S. INDUSTRIAL CONTROL FIRM BRANCH MANAGERS



N = 98 branch managers, in two samples.

Economic value added by BM competencies =.80 x .34 x $1.69 million = $459,680

As can be seen from Figure 3, four competencies account for 80% of variance in BM economic performance. Previous years branch revenues were included in the regression equation to test the usual objection that "getting a good district" (fast growing local economy, weak competition, etc.) really accounted for performance. (Having a "good branch" the prior year made some difference, 10%, but far less than the 80% due to competence. The competency model predicts .80 x .34 x $1.69 million = $459,680 in economic value added.

Figure 3 also suggests some of the ways competencies combine in practice. Initiative by itself accounts for 48% of performance, but Initiative to do what? Initiative correlates .73 with Achievement, and .34 with Team Leadership, so superior BMs probably spend most of their time improving performance and communicating their vision and goals to subordinates.

Organizing Competency Data in Competency Models

> Anthropologists asked a primitive South Sea Islanders to organize common tools and utensils in "logical" order. The Islanders placed a hunting spear on the ground, then the carcass of a pig, a skinning knife, rocks for a fire ring, wood, a large cooking pot, plantain leaves used for plates, and small knives used for manipulating bite-sized food. The anthropologists then asked the Islanders to organize the tools and utensils the way a "stupid" person would. The Islanders put all the spears together, all the knives together, all the cooking pots together …

Development of competency models has evolved from

1. Lists of nice-to-have characteristics: hundreds of behavioral indicator "atoms"

2. Clusters of similar behavioral indicators in competency "elements" (the Aristotelian mutually exclusive, collectively exhaustive taxonomy the anthropologists expected),

3. Combinations or "molecules" of competencies used together to accomplish a task or job effectively,

5. Specific "molecular" combinations of competencies which fit specific "receptor sites" tasks or jobs, and

6. "Dynamic" competency models which describe competency combinations used in sequential steps of accomplishing a task or dealing with a critical situation (the model most logical to the Islanders).

- **Competency "molecules"** are combinations of competency elements which act together to produce effective performance in task situation.

Competency model lists of behavioral indicators or competency elements do not provide any information about how these elements combine or act together to produce effective behavior. Molecules have very different properties than their constituent elements e.g.

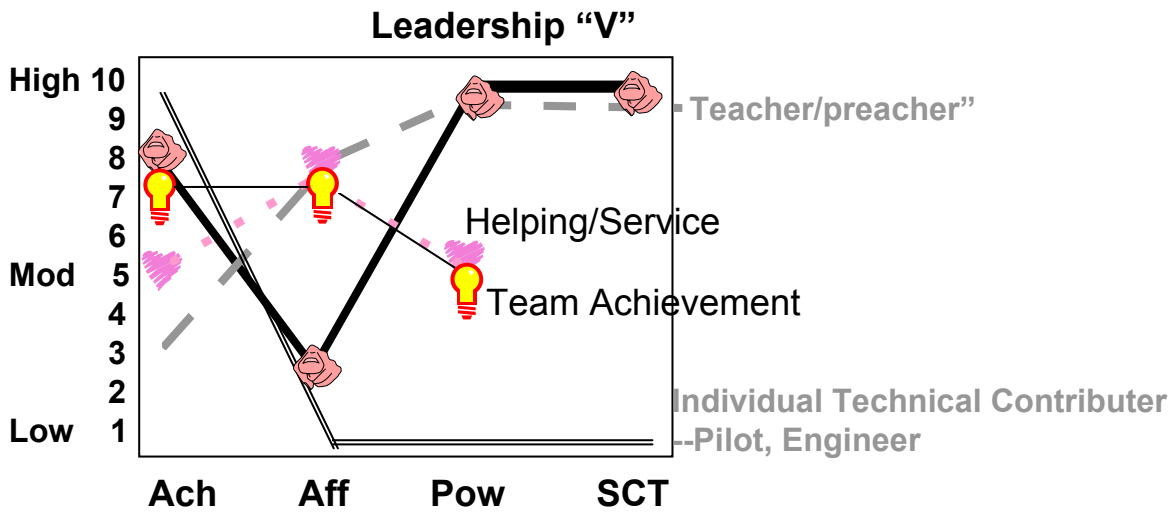
the liquid H_2O, water, differs completely from its constituent elements, the gases hydrogen, H_2, and oxygen, O_2.. "Laundry list" competency models are similar to biological analyses which define a human being as being composed of hydrogen, oxygen, carbon and nitrogen. This "model" *might* distinguish a human from a rock but would not distinguish a human from a tree.

Motive profile "molecules" have been known to predict behavior in different job families for many years. For example, the "leadership motive profile" (McClelland & Boyatzis,1982) : moderate-high achievement motivation (n Ach), low- moderate affiliation motivation (n Aff), high power motivation (n Pow) and high self-control (SCT) is a four-element competency "molecule" which predicts success in leadership and management jobs:

$$\underline{n}\ \text{Ach}_8\ \underline{n}\ \text{Aff}_3\ \underline{n}\ \text{Pow}_9\ \text{SCT}_{>5}$$

where subscripts are normal distribution decile scores e.g. 3=30%ile, 9=90%ile as shown in Figure 4..

Figure 4. Motive Profile "Molecules"



Achievement: do better against results standards
Affiliation: make friends, express caring for others
Power: impact, influence others
***SCT*:** self control ("socialized" v. "personalized" Power)

Helping/Service competencies are driven "helping" motive profile, a molecule of moderate-to-high Affiliation with moderate Achievement and moderate Power motivation (Kolb & Boyatzis, 1970; McClelland, 1978)

$$\underline{n}Ach_5\underline{n}Aff_7\underline{n}Pow_5$$

Competency studies of high performing members in effective self-managing workgroups have identified a new "molecule" called the "team achievement" motive profile, composed of moderate-high achievement, moderate-high affiliation and moderate power motives.

$$\underline{n}Ach_8\underline{n}Aff_8\underline{n}Pow_5$$

A worker in a toy manufacturer captured the essence of this findings as "we are all good friends, having lots of fun together making more and more widgets, better and better, faster and faster…we can influence each other when needed, but don't have to play a lot of politics." "Fast, focused, flat, friendly, self-confident, fun" teams are increasingly common in entrepreneurial high tech organizations where fast product development cycle times are essential to meet competitive pressures

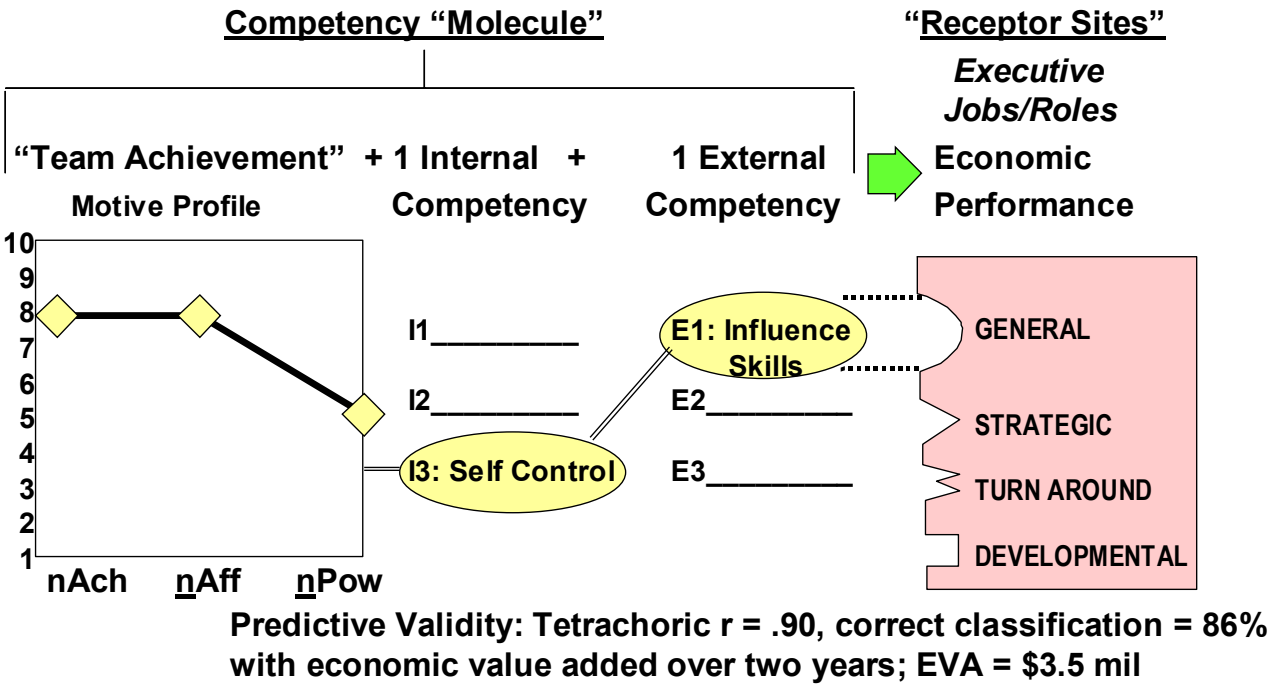
- **Molecules X Receptor Sites** (specific jobs, tasks)

  All competency models imply a job ("receptor site")-person (competencies) match assumption*: the better the fit between the requirements of a job and competencies of a person, the higher will be the persons' job performance and job satisfaction*. (Caldwell, 1991). This assumption is infrequently tested empirically because valid competency models for jobs and accurate competency assessments of people are rare.

  Competency models based on identifying the specific competency "molecules" to specific job/role receptor sites have shown criterion validities as high as .8. For example, As shown in Figure 5, the "team achievement" motive profile, plus one of three internal (e.g. self confidence) and one of three external (e.g. organizational awareness) competencies predicted economic outcomes over two years at r=.90 for four types of executives (strategic, turnaround, developer and general) executives in Pepsico (McClelland, 1998) . Molecule—receptor site

matching was key to getting this result e.g. "internal" competencies of achievement motivation and conceptual thinking predicted success for strategic executives but not more "hands on" turnaround and developer managers.

Figure 5. Competency Molecules which predict Superior Economic Performance over two years in Pepsico



Ref. McClelland, D.C. (1998). Identifying competencies with the behavioral event interviews. Psychological Science. Vol. 9, No. 5, Sept

Dynamic Competency Models

Recent competency (and neuroscience) research indicates that competence is always a "molecule" or combination of competencies, usually with four components

(1) *Motivation*: one or more *operant* motives and one or more *respondent* values, preferences or attributions;

(2) *Observation:* information seeking, interpersonal understanding or organizational awareness;

(3) *Understanding:* cognition: declarative knowledge content "data" bases, conceptual "rule" bases, and analytic or conceptual processing (IQ); and

(4) *Decision :whether* to act, a product of emotional intelligence "gut feel" and rational business case expected value analysis; and *how* to act, rational planning, logistics, etc.

(5) *Action:* behavioral skills.

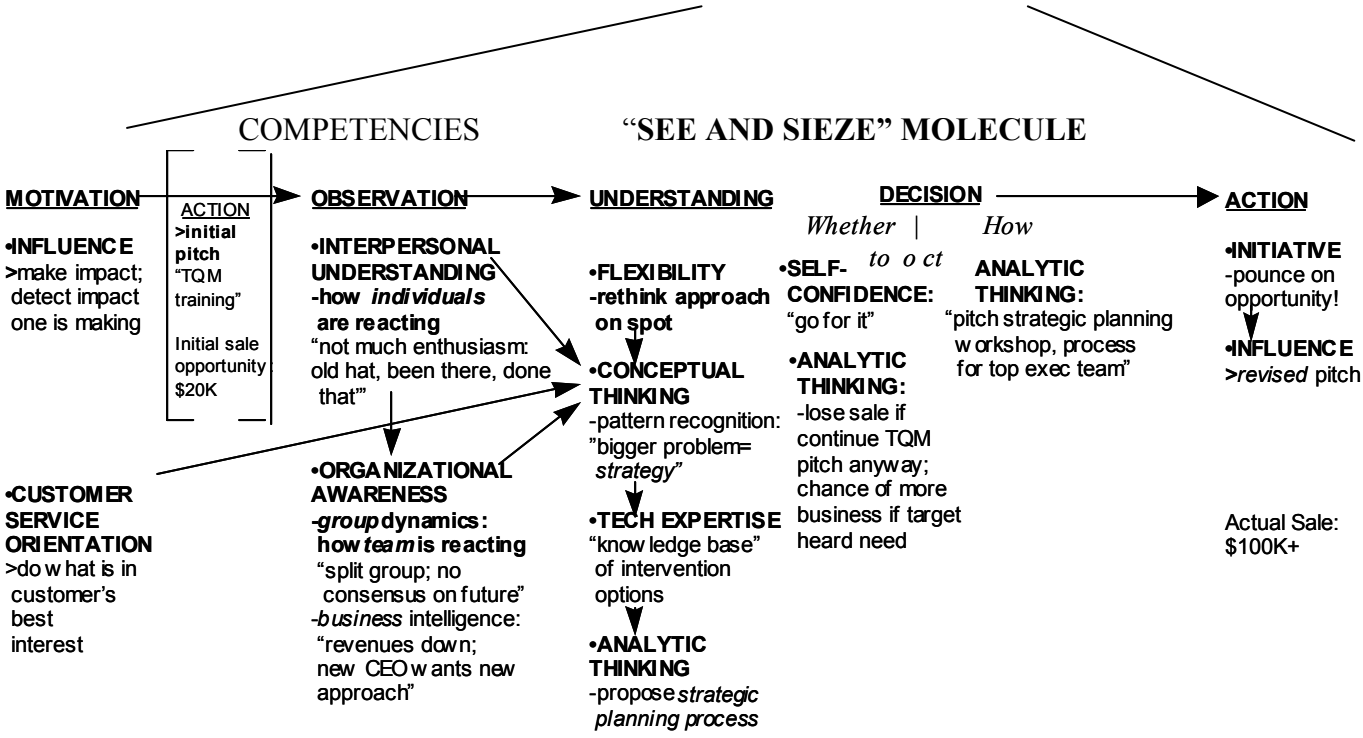
For example, effective influence in sales is a "motivation-observation-understanding-decision-action" (MOUDA) molecule composed of eleven competencies is shown in Figure 6. Driven by motives to Influence and add value (Customer Service Orientation), superior salespeople seek information using Interpersonal Understanding and Organization Awareness competencies; identify client needs using Conceptual Thinking, and then initiate to persuade prospects using Influence Skills.

Interpersonal Understanding is a prerequisite for effective influence and persuasion. An understanding of the unspoken feelings, concerns and agendas of those individuals, and the informal "political" alliances, rivalries, key decision-

makers, etc., of organizations, is needed before one can effectively influence individuals or organizations.

Figure 6. Competency Molecules mapped to a Step and Critical Situation in a Consultative Selling Job

STEP IN PROCESS	4 Call client contact to "scout" personalities, issues likely to come up in meeting	**5** **Initial meeting with client management buying decision makers**	6 Call contact to get feedback on meeting, tips on points to make (or not make) in proposal
CRITICAL TASK SITUATIONS		**Audience not buying initial pitch; other needs offer different selling opportunities**	



In Berger, L. A. & D. R Berger Eds (2003) *The Talent Management Handbook*, New York NY: McGraw-Hill

VALIDATION OF COMPETENCY MODELS

Competency models are validated in three ways:

- concurrent validation, (meaning competencies measured currently correlate with current results,
- cross validation (scientific replicability), meaning competencies which predict success and differentiate in one sample also predict success and differentiate in another sample of superior vs. average performers; and
- predictive validity, a much sterner and more valuable standard meaning competencies measured at time 1 actually predict economic performance at a future time.

Predictive validity is the obvious objective and challenge for talent management: candidates hired or fast-tracked now in fact produce superior results in the future

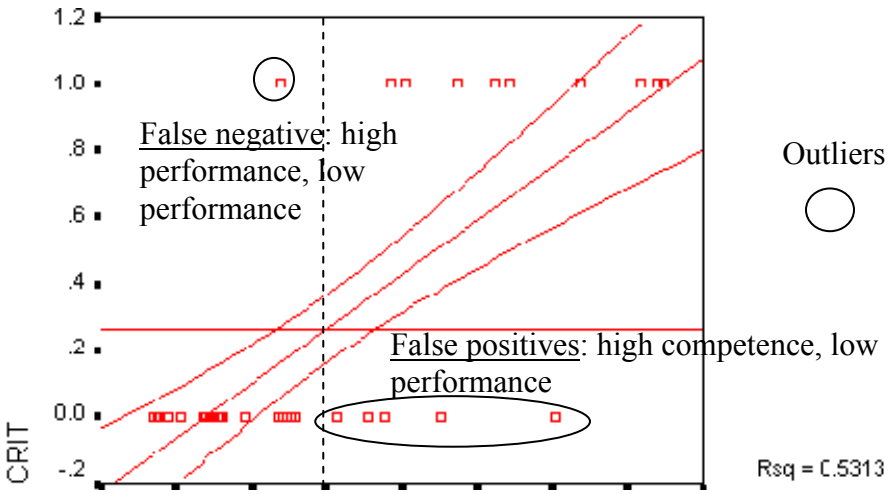
CASE STUDY

Incon cross validated Branch Managers competencies in two Incon divisions (Figure 7).These divisions had quite different BM revenue and profit data so absolute economic data are not comparable. The regression graph shows standardized competency scores predicts superior performers in both divisions (r^2=.53). Outliers include one "false negative" BM who achieved +1SD economic performance despite below average competencies, and five "false positives" who scored above average in competence but achieved below average profitability. Competency scores correctly classified 80% of superior and average performers.

The cross validation study validated the four competencies which predicted superior economic performance in the first division study ACH (achievement), INT

(initiative), TL (team leadership) and TW (teamwork). Five additional "predictor" competences were also identified by the model: SCF (self confidence). DEV (developing others), CO (concern for order and quality) and CSO (customer service orientation).

FIGURE 7. CROSS VALIDATION OF COMPETENCY MODEL PREDICTION FOR BRANCH MANAGERS IN TWO INCON DIVISIONS



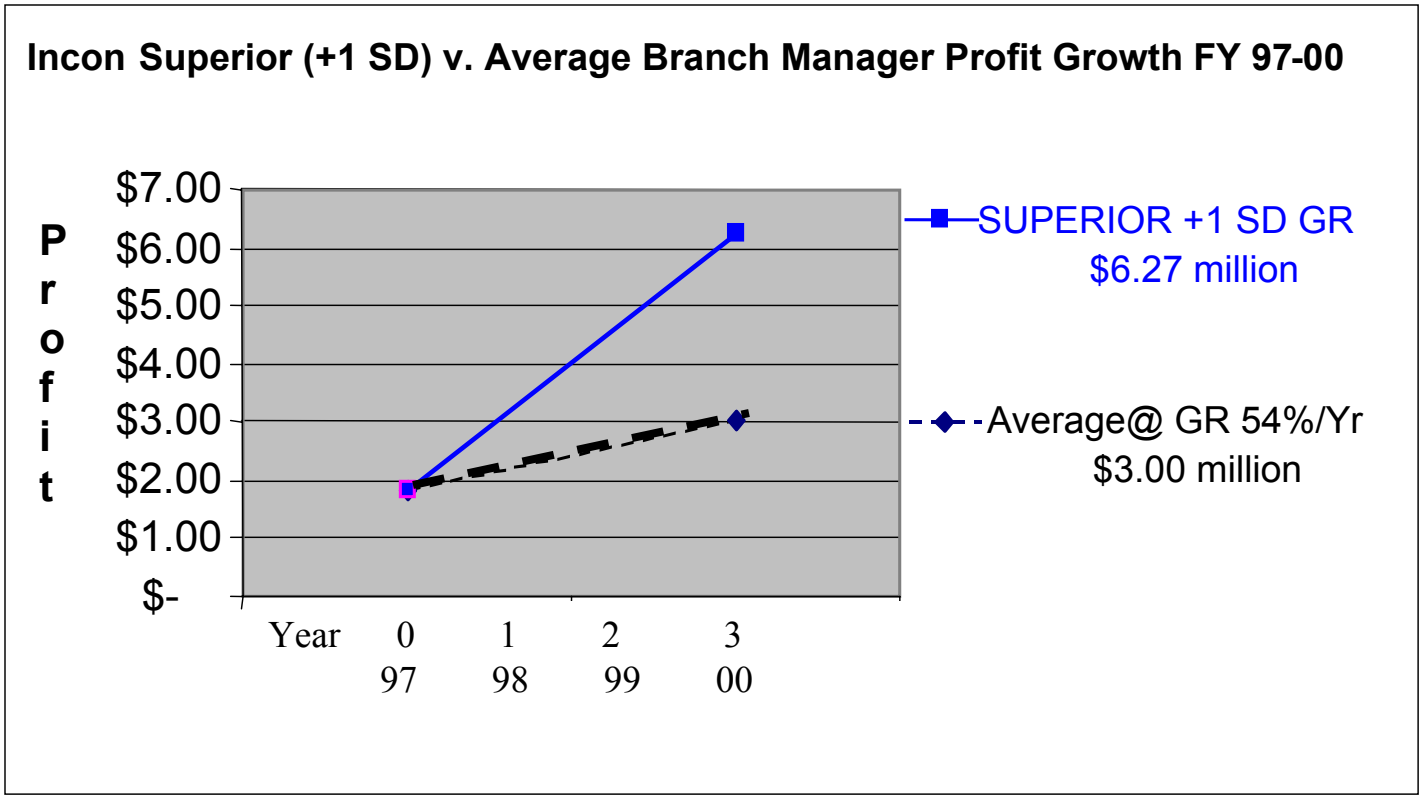
Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change
1	.729[a]	.531	.402	.35	.531

a. Predictors: (Constant), SCF, DEV, TW, TL, CO, INIT, CSO, ACH

Figure 8 shows the predictive validity of BM competencies measured in 1997 to profits in the next three years 1998-2000. BMs +1 standard deviation above the mean generate more than twice (113%) profit growth than average BMs (54%), a difference of $3.27 million over the three years.

Figure 8. Three year Predictive Validity: Incon Branch Manager Competencies measured in 1997 prediction of profits in years 1998-2000



Each superior Incon Branch Manager is worth an additional $3.27million to the firm over 3 years, $1.09 million per year. A valid competency model costing $100,000 which selects even one superior BM will return 1000% over 1 year.

APPLICATIONS

Detailed discussion of applications of competency models are beyond the scope of this chapter, but validated models have great power to help HR professionals build business cases by estimating the potential economic value added from selection, feedback, training, and performance management.

Once the following variables are known

- economic value of superior performance (1 effect size = +1 standard deviation above the mean) from a criterion sample of job incumbents
- the percentage effect size differences HR applications can make (from published meta-analytic studies -- see Spencer, 2000);

  and
- the number of employees affected by the HR intervention.

  the equation for potential economic value added is $EVA of 1 effect size X average %improvement in effect size HR application can make X the number of employees impacted by the HR application, as shown in Table 6.

Table 6. Calculation of Expected Economic Value Added by HR Applications

HR Intervention	$EVA of 1 effect size (millions)	Average %effect size HR application can make	#employees impacted by HR application	Total HR Economic Value Added (EVA)
Selection:	$1.000	19%	20	$3.800
Feedback	$1.000	11%	20	$2.200
Training:	$1.000	40%	20	$8.000
Performance Management	$1.000	60%	20	$1.200

CASE EXAMPLE

Incon calculated the expected EVA and ROI from its pilot BM training program using a .44 expected effect size shift in average trainee performance. The expected value added (EVA) per trainee is .44 x $456,300.= $200,772, and total return from training 28 BMs at a cost of $8,000 per trainee, $5,621,616, a potential 2,410 percent ROI, as shown in Table7.

If the firm's cost of capital is 8.5 percent, and the standard deviation of effect size shifts from training is .27, the effect size shift needed to achieve an adequate return, and the probability of successfully achieving this return, can he calculated as follows:

$$es_{ROI} = \frac{I(1+\%CC)}{VA_{es}\,N} = \frac{\$224,000*(1+.085)}{\$200,772*28} = .04$$

Where

es_{ROI} is the effect size needed to achieve the required ROI when

I = investment in training ($224,000),

%CC = the firm's cost of capital (8.5%)

VA_{es} = value added ($456,300) by the expected es (.44) per trainee = $200,772, and

N = number of persons trained (28).

TABLE 7 BUSINESS CASE FOR TRAINING.

ES shift from training	0.44
EVA of training per person	$456,300
N trained in U.S.	28
EVA from N trained	$5,621,616
Investment in U.S. training	$224,000
ROI	2,410%

Incon further calculated a "sensitivity analysis" to check the minimum effect size shift needed to justify the investment in training, and the probability of achieving this effect size shift and return.

Probability of success = p @ z

Where

$$z = \frac{\mu es - es_{ROI}}{SD_{es}} = \frac{(.44) - (.04)}{(.27)} = 93\%$$

μ es = mean effect size expected from training, .44 (Spencer, 2001)

es_{ROI} = effect size needed for an ROI equal to the firms cost of capital (calculated above)

SD_{es} = standard deviation in es from training (from Spencer, 2001)

p =the probability at the calculated z value (available from any statistics text).

In this case, an effect shift of 4 percent justifies the investment in training, and the probability of achieving an acceptable return is 93 percent, as shown in Table 8. The business case for training is reasonable.

TABLE 8. SENSITIVITY AND SUCCESS PROBABLILITY ANALYSIS

ROI required	8.50%
ES required for desired ROI	0.04
z @ ES shift required	1.481
p success @ z	93%

STEP 7. EVALUATION

The truest test of a competency model is whether its application actually adds economic value to the organization that invested in it. If competencies identified, assessed, trained, given feedback on, and used as the basis of goal setting were valid, the application program should produce a significant return on investment.

CASE EXAMPLE

Incon evaluated the change *in* trainees' competence and calculated the economic value added in comparison with the competence and EVA of the control group.

As shown in Table 9, the competency research, training, feedback, and goal-setting intervention at Incon appears to have significantly increased participant branch managers' sales and profits, producing a 613 percent ROI. Trainees' return on sales decreased (insignificantly) compared to that of the control

group—perhaps because trained managers were investing in revenue-increasing marketing and area expansion efforts. However, trainees' increased revenues more than made up for this decline. The 0.04 effect size shift achieved by training was only 10 percent of the expected 0.44. This shows that even a very small shift in performance can result in significant statistical and economic results when the economic value of the problem or opportunity in the business case is large.

TABLE 9. TREATMENT GROUP VERSUS CONTROL GROUP PERFORMANCE OVER ONE YEAR AFTER TRAINING.

	Revenue	Operating Income	Profit
Trained group *(N=23)*	$3.117M	0.3%	$249,000
Control group *(N* = 7)	$1.660M	0.7%	$192,000
Difference	$1.457M	- 0.4%	$ 57,000
p (t- test)	*< .04*	*n.s.*	*< .02*

- Es shift from training: $57K/$456K ~.125
- Investment: $8,000/BM trained x 28 BMs trained = $224K
- Return: +$57K Profit/BM trained x 28 BMs trained = $1,596K additional profit
- ROI = 613 percent

Trainees increased revenues and profits significantly more than the control group. Trainee versus control return on sales did not differ significantly. The .125 effect size shift achieved by training and feedback is similar to that reported by McClelland (1998) for COMPETENCY assessment feedback to executives. That the .125 effect size is 28 percent of the .44 meta-analytic mean for all training

programs suggests that more in-depth training involving action learning projects could increase return on training investments.

SUMMARY

The future of competency modeling will be increasingly precise definition of competencies (likely at the neuroscience level) and competency molecules which predict economic value added in precisely defined jobs/roles. Better quantitative tools more reliably measure competencies and predict business The "junk science" of flip chart models" lacking reliability and validity will be replaced by models grounded in science.

BIBLIOGRAPHY

American Educational Research Association, American Psychological Association, National Council On Measurement In Education (1996) Standards For Educational And Psychological Testing. Washington DC: American Psychological Association.

Bakke vs. Regents of the University of California (1978) 17 FEB.

Barrett, G.V. (2000, April). *Emotional Intelligence: The Madison Avenue approach to professional practice*. Paper presented at the 15th annual meeting of the Society for Industrial Organizational Psychology. New Orleans. LA.

Boyatzis, R.E. (1982). *The Competent Manager*. New York, NY: John Wiley

Boyatzis, R.E., S.S. Cowen & D.A.Kolb (1995) *Innovation in Higher Education*. San Francisco, CA: Jossey-Bass.

Burke, M. & Day, R. (1986). A Cumulative Study of the Effectiveness of Managerial Training. *Journal of Applied Psychology*.

Falcone, A.J., J.E. Edwards and R.R.Day. (1986) *Meta-analysis of personnel training techniques for three populations* . Paper presented at the Annual Meeting of the Academy of Management. Chicago, IL.

Morrow, C., M.Q. Jarrett, and M. T. Rupinski (1997). An investigation of the effect and economic utility of corporate-wide training. Personal Psychology. 1997. **50.**

Damasio, A. (1994) Descartes' Error: emotion, reason and the human brain. New York, NY: Grossett/Putnam.

Goleman, D. (1997) Generic Emotional Intelligence Competence Framework. Trenton, NJ: Rutgers Univ. Fetzer Consortium

Hunter, Schmidt and Judiesch (1990)). Individual differences in output variability as a function of job complexity. *Journal of Applied Psychology, 75(1),* 28-42.

Jones, C. (1986). Programming Productivity. New York, NY: McGraw Hill.

Jones, C. (1991). Applied Software Measurement. New York, NY: McGraw Hill

Latham G. E. & K.N.Wexley, K. N. (1981) *Increasing Productivity Through Performance Appraisal*. Reading, MA: Addison-Wesley Publishing Co.

Latham, G.P. & Locke, E. A. 1979) Goal setting: a motivational technique which works. *Organizational Dynamics*, 68-80

Martin, J. (1990). *Rapid Application Development*. New York, NY: McMillian

McClelland DC (1998). Identifying competencies with the behavioral event interviews. *Psychological Science*. Vol. 9, No. 5, Sept.

Miron & McClelland (1979). The Impact of Achievement Motivation Training on Small Business. *California Management Review*. XXI No 4.

Rondina, P. (1988, October). Impact of competency-based recruiting techniques on dropout rates in sales training programs. Paper presented at the McBer 25th Anniversary Symposium, Boston, MA.

Spencer, L. and C. Morrow, (1998) Calculating the ROI on Competence. London, UK: Linkage Competency Conference Seminar. 2 November 1998.

Spencer, L.M. & Spencer, S. (1993). *Competence at Work*. New York, NY: John Wiley & Sons.

Spencer, L.M. (1986). *Calculating human resource costs and benefits.* New York, NY:John Wiley & Sons.

Spencer, L.M. (1997). Project Management Competencies. San Diego, CA. Paper presented at Engineering Construction and Contracting (EEC) Association Annual Conference. (5 Sept. 1997). Available from author.

Wiener, R.L. (1999) Extending Daubert beyond scientific expert testimony: Technical and other specialized knowledge must be reliable to be admissible. *Psychological Monito*r Washington DC, American psychological association. VOLUME 30 , NUMBER 6 June 1999