PERSONNEL PSYCHOLOGY
1988, 41

SIMULATION-BASED ASSESSMENT OF MANAGERIAL COMPETENCE: RELIABILITY AND VALIDITY

SIEGFRIED STREUFERT, ROSANNE POGASH, MARY PIASECKI
College of Medicine
The Pennsylvania State University

Prior work based on complexity theory has attempted, with some success, to predict general and managerial performance in complex, uncertain, and fluid task settings. The present paper evaluates a quasi-experimental simulation technique that was specifically designed to measure the impact of individual differences in a number of managerial styles (including a style reflecting cognitive complexity) on executive performance. Twenty simulation-based measures were tested for reliability and validity. On the basis of the data obtained from two separate samples, it was concluded that this quasi-experimental simulation technology may be useful for assessing a number of managerial styles that are not currently tapped by other measurement methods. Research results reported in this paper, as well as results obtained in other concurrent efforts, are summarized.

For more than two decades, researchers have employed constructs and measurement techniques of complexity theory to test performance in complex perceptual and decision-making tasks (cf. reviews by Scott, Osgood, & Peterson, 1979; Streufert & Streufert, 1978; Streufert and Swezey, 1986). In general, the application of complexity theory to individual differences in human task performance has been quite successful. For example, it has been shown that more cognitively complex individuals exceed less complex counterparts in interpersonal perceptual accuracy (Streufert & Driver, 1967), communication skills (Hale, 1980), and the capacity to discern an opponent's strategy (Streufert & Driver, 1965).

Other research indicates that cognitively complex individuals differ from those with lesser complexity in the content and flexibility of their attitudes (e.g., Bhutani, 1977; Linville & Jones, 1980; Mizerski, 1978; Streufert, 1966), attributions of causality and responsibility (Streufert & Streufert, 1969), information search and information utilization (Hendrick, 1979; Karlins & Lamm, 1967; Streufert, Suedfeld, & Driver, 1965), and related characteristics. Obtained individual differences become especially evident where task load or other forms of task- or situation-related stressors tend to be intermediate (i.e., optimal; cf. Streufert, 1970, 1978; Streufert & Streufert, 1978). In contrast, excessive task loads or deprivation may

Correspondence and requests for reprints should be addressed to Dr. Siegfried Streufert, Department of Behavioral Science, Pennsylvania State University College of Medicine, Hershey, PA 17033.

negate benefits that a person might gain from the capacity to think and perform in a cognitively complex fashion (cf. Streufert, 1970; Suedfeld, 1964a, 1964b, 1978).

Early complexity theorists and researchers have generally investigated theoretical propositions in the laboratory, enrolling primarily college sophomores as subjects. Success in predicting student behavior on the basis of complexity theory has encouraged others with applied interests to expand theoretical predictions to managerial behavior in the private and public sectors. Their primary focus has been on individual differences in persons with considerable decision-making responsibility (e.g., Schroder, 1982; Streufert & Swezey, 1986; Suedfeld, Corteen, & McCormick, 1985; Suedfeld & Rank, 1976; Suedfeld & Tetlock, 1977). As a consequence, researchers have found it useful to develop additional measurement methodologies that may be relevant to the impact of cognitive complexity on a variety of managerial activities.

Because of the interest in effects of cognitive complexity on managerial activities, two previously recognized measurement problems have become more evident. First, a person's cognitive complexity reflects a *style* that is rather difficult to assess with objective paper-and-pencil instruments: Complexity is concerned with "how" persons think and behave. While people are generally quite aware of "what" they think or do, they find it very difficult to conceptualize "how" they think or how they arrive at decisions. Responses to an assessment of stylistic (cognitive) complexity, however, should reflect the "how" of a person's thoughts and actions. Because of an inability to conceptualize the "how" of thinking processes, a person's responses to relevant paper-and-pencil queries may be irrelevant and/or may be seriously confounded with social desirability.

A second measurement problem can be equally problematic. Cognitive complexity represents only *one* of a constellation of managerial information-processing and decision-making styles. Various styles may jointly or interactively affect managerial competence and success (cf. Streufert & Swezey, 1986). To predict managerial effectiveness, it is necessary to assess and evaluate several styles simultaneously. Moreover, such an assessment should ideally occur in a managerial (or equivalent) setting, yet it should not be biased by task-specific knowledge and/or experience that a manager may possess—or may have gained via training.

Further, assessment should remain independent of the *content* of executive thought and action (cf. Streufert & Swezey, 1986). Content of thought reflects such things as attitudes, beliefs, and specific action tendencies that might be based on beliefs, prior knowledge and/or task-relevant experience. Of course, attitudes, beliefs, and other content components of "what" executives think or do can be of considerable importance for the achievement of managerial success.

Similarly, motivation levels play an important role in executive functioning. However, variables of this nature have already been adequately assessed with a variety of existing techniques. Moreover, assessment of thought and action content is likely of decreasing value as executives advance to more senior levels or as managers enter tasks or jobs that are associated with novelty, complexity, uncertainty, and fluidity (cf. Streufert, 1986). Lack of motivation, inadequate intellectual capacity, inappropriate or insufficient experience, ineffective attitudes, and other undesirable "content" characteristics of executive functioning tend to be recognized relatively early in executive careers and, typically, result in stagnation of advancement or even in loss of employment.

Simulation technology may be the ideal and, potentially, the only viable method of simultaneously assessing several untapped or neglected executive styles. However, the typically employed "free" simulations (cf. definitions by Fromkin & Streufert, 1976) are inadequate for this purpose. Participants in free simulations make decisions that can drastically modify their subsequent task environment. As a result, comparisons among different individuals (or groups) *after* they have participated in a free simulation for some amount of time can be difficult or impossible. To correct this deficiency, quasi-experimental simulation technology (cf. Streufert & Swezey, 1985) was chosen by the present authors to permit comparison among diverse participants and to allow the calculation of meaningful reliability and validity coefficients.

Quasi-experimental simulations maintain much of the realism and all of the potential complexity, fluidity, and uncertainty that can be programmed into free-simulation techniques. Also, the quasi-experimental simulation permits experimenter (assessor or trainer) control over information flow and content, as well as other task characteristics, events, and demands. As a result, comparison among individuals (or groups) who have participated in the same simulation at different times becomes possible. Comparison of performance with an established criterion can be achieved. However, it might be argued that the introduction of "programmed" (i.e., fixed) events might diminish reliability and validity. Before such a simulation can be used for assessment and/or training purposes, data on reliability and validity should be obtained. This paper reports on those data.

Simulation Design

An underlying software program that operates simulated events and simultaneously collects performance data—and later analyzes it—was developed. The simulation functions as a man-machine system, with the participant's "assistant" being an operator of an enhanced IBM AT computer. The computer provides the participant with continuous visual and hard-copy information about ongoing events. Participants tend to believe that they have direct, even though sometimes delayed, impact upon those events. In fact, however, one-half of the simulated events are fixed and presented to all participants at exactly the same time point in the simulation. The remaining events are either partly responsive (with pre-specified endings to actions of participants) or, if insufficient decisions are made by a participant, selected from random (but not especially important) stored messages. Each participant receives the same quantity of information at the same fixed points in simulated time.

During the evening prior to simulation participation, participants spend several hours reading detailed, informative manuals about the next day's task. Two lengthy manuals contain considerable detail about the task. Task and setting are introduced with additional video information (30 minutes) the following morning. The simulation itself lasts six hours. The specific task *content* on which simulation scenarios are based is unfamiliar to participants, even though they would recognize that such tasks do exist. These computer-assisted simulation designs are similar to earlier, manual quasi-experimental simulations by Streufert and associates, which—despite their cumbersome nature—produced a wealth of data (cf. Streufert & Streufert, 1978, and Streufert & Swezey, 1986, for reviews).

Two simulation scenarios are available. In one, the participant is appointed "disaster control coordinator" in charge of various public and civic organizations of "Woodline County" at the time of a potential emergency. In the second scenario, the participant assumes the position of temporary "governor" of "Shamba," a developing country plagued by internal unrest, hostile neighbors, an ineffective and untrustworthy military establishment, and serious economic woes.

While the two simulation scenarios differ greatly in content (to avoid unintended training of those participants who are sequentially exposed to both scenarios), demands made on participants' managerial functioning are designed to be highly similar. For example, experienced information load, success, and failure levels are identical. A serious emergency requiring quick and decisive action occurs at the same time point in both scenarios. The nature of the emergency, however, differs. In the "Woodline County" simulation, a dam break destroys several villages in a canyon. In the "Shamba" simulation, rebel forces capture and temporarily hold a strategic mid-sized town.

The two simulation scenarios are designed as parallel, novel, complex, fluid task environments where action outcomes are, in part, uncertain. The scenarios present a managerial task environment that is best dealt with via a number of diverse managerial activities, including preventive action, use of strategy, planning, responsive action, information search, timeliness of responsive action, use of opportunism, and more.

Once the simulation begins, participants may engage in any action permitted by their resources. An experimenter "assistant" of the participant records in the computer system all decisions, future plans, prior related actions, and information that led to any action. Information on current events (either preprogrammed or partially responsive) is periodically provided to the participant via a video screen and in hard-copy format. Information flow is a function of predetermined load levels.

After six one-hour task periods, the simulation is terminated. To avoid potentially confounding "end effects," participants are not informed about this end point. The computer program subsequently provides hard copy of some 40 performance measures. Some of these measures are identical or similar to measures that were previously employed in research by Streufert and associates (cf. the measures listed in the Appendix of Streufert & Swezey, 1986). Many of those earlier measures had been developed on the basis of complexity theory predictions.

Both similarities and differences between the simulation technology and assessment center technology exist. Of course, assessment centers differ from each other; in other words, specific comparisons apply in many but not all cases. While most assessment centers employ multiple measurement techniques, the present simulation technology involves a single lengthy and realistic task, which generates multiple performance measures. The present simulation also differs from most other simulation methods that are employed by assessment centers. While the latter generally represent "free simulations" (see above), the present technique is quasi-experimental in its control of event input to participants.

While many assessment center techniques focus primarily on content (specific knowledge, experience, etc.) and interpersonal functioning, many, but not all, of the simulation demands and measures focus on structural style (i.e., the "how" of managerial functioning). There is no group discussion among assessors (cf. Byham & Thornton, 1986) that might bias or modify assessment outcome (Heriot, Chalmers, & Wingrove, 1985; Russell, 1985). Bias introduced by individual rater judgment is also eliminated: performance analysis is attained via a computer program. Group member contributions to the task environment are simulated, eliminating potentially uncontrolled effects of other persons on the actions and performance of a participant. Of course, the latter characteristic limits direct assessment of a participant's interpersonal effectiveness.

The differences and, to some extent, advantages of the simulation technique do not imply that the present technology is always better than assessment center procedures. However, it appears uniquely qualified to assess certain neglected aspects of managerial performance that are not typically covered by most assessment center technologies, especially performance aspects involving structural style of thought and action.

The present paper reports on reliability and validity of 20 selected simulation measures. Since reliability data and validity data were obtained from different groups of participants at different times, the results will be reported in separate sections.

Reliability

Method

Eighty-one adults with managerial experience participated as individuals in two simulations. The scenarios were presented in random order. Performance scores for each participant were calculated. Scores can be separated into three groups: measures of content (i.e., *what* the decision maker did), measures of structural style (*how* the decision maker approached problems at hand), and mixed measures. Since content measures are primarily sensitive to scenario specific characteristics, high intersimulation reliability (intercorrelations of performance scores for the two simulations) was not expected. Consequently, content-oriented measures were not included in this analysis. Variable or moderate correlations might be expected for mixed measures that are partly stylistic and partly content oriented. Meaningfully significant intercorrelations should be expected where measures reflect structural styles.

This paper focuses primarily on 12 *structural* measures of management style that were calculated by the simulation system. The measures selected for this purpose loaded highly on factors previously extracted from performance scores (cf. Streufert, Pogash, Piansecki, Repman, & Swezey, 1986). For comparison, eight mixed measures of activity or speed of action were also tested for reliability and validity. Each of the resulting 20 measures is considered below:

Structural Measures of Managerial Style

Measures of managerial style, concerned primarily with how decision makers approached problems at hand were :[1]

1. *Diversity of Action (number of decision categories).* This measure is somewhat akin to the concept of cognitive differentiation (a component of cognitive complexity) and reflects the degree to which a number of diverse (independent dimensional) actions are represented among the actions taken. For example, if a participant separately orders several investments in durable goods, of whatever nature, the actions would be placed within

[1]The name designation for those measures that appear in software-based hard copy of earlier simulation results represent technical terminology assigned at the time of software development. For comparison purposes, those name designations are provided in parentheses.

the same decision category. On the other hand, investments, sales of manufacturing facilities, airport construction, or requests for outside aid would all represent different decision categories. The simulation environment is sufficiently rich to allow ample diversity of action (as well as ample activity on all other measures discussed below). Limits for diversity of action are set only by the number of individual actions upon which a participant decides.

2. *Use of Strategy (forward integrations).* This measure is a basic indicator of cognitive integration (another aspect of cognitive complexity) and measures the degree to which a person engages in actions that provide the foundation for subsequent different actions. For example, if a disaster control coordinator in the Woodline County simulation arranges for buses to wait at local schools to speed potentially needed future evacuations and, when an emergency actually threatens, orders those buses to take the school children to high ground, he or she is credited with one unit of use of strategy. Potential use of strategy, again, is restricted only by the number of actions upon which a participant decides.

3. *Elapsed Time Between Strategic Actions (time weight).* The length over time across which any single plan spans is assessed. This measure is concerned with the average length of time a person uses to prepare future actions. For example, if, during a potentially threatening emergency, buses were ordered to schools one-half simulation hours prior to the time they had to be put to use, a score of .5 would be obtained. If, on the other hand, the buses had been placed on locations two simulation hours earlier, a score of 2 would have been generated. The score for this measure represents the mean time length of all integrations achieved during a simulation.

4. *Utilizing Opportunity (number of backward integrations).* This measure considers whether a decision maker utilizes previous actions and events to aid in the execution of effective current actions. This capacity is viewed as a "lower-order" style by some complexity theorists. For example, if the emergency control coordinator had parked some buses near one of the schools for a maintenance check and, upon notification of an immediate emergency threat, used those vehicles rather than others located farther away, he/she would receive one credit for utilizing opportunity. The potential for utilization of opportunity is restricted only by the number of actions upon which a participant decides.

5. *High-Level Strategic Planning (quality of integrated strategies: "QIS" or "weighted QIS").* The degree to which multiple stepwise strategies are interrelated across past, present, and future planning is assessed. In the "Use of Strategy" measure we are concerned with any single strategic interrelationship between two actions; the present measure counts the number of *additional* strategic actions that directly precede the initial decision point of a strategic sequence. In addition, the measure counts *additional*

strategic actions that directly follow from the end point of a strategic sequence. In other words, the measure assesses the complexity of a strategic sequence as well as the number of steps involved in that sequence. For example, if a disaster control coordinator, anticipating potential problems, had alerted drivers to be available all day; had, as a serious problem became more likely, ordered buses and drivers to remain at the schools; and upon an imminent threat of emergency, had ordered the drivers to transport children to an emergency shelter—in the meantime planning to use the buses to transport food and cots to that shelter after the children had been delivered—and had then carried out this last intention as well, one credit would be given for a prior integration (alerting drivers) and one for a subsequent integration (using the buses to obtain cots and food from some other location). The total number of prior and subsequent integrations associated with all credited scores for use of strategy during the simulation are added. Since the capacity for high-level strategic planning is not normally distributed, a logarithmic transformation of the obtained value (to eliminate curvilinearity) is employed as a person's score. The potential score is restricted only by the degree to which a person has engaged in use of strategy.

6. *General Planning (planned integrations).* Plans for subsequent actions are, at times, not completed because multiple strategies may render them superfluous. This measure, in contrast to other measures of strategy, considers the frequency of such contingent planning. It counts the number of times a person utilizes preparatory action to generate the preconditions for a later action that is not actually carried out. For example, if an emergency control coordinator had alerted yet another bus depot in case intended buses were cut off by landslides or high water—an alternate procedure that need not be activated—credit for a planned integration would be given. Potential scores for general planning are restricted only by the number of actions upon which a participant decides.

7. *Strategic Actions within Groups of Coherent Activities (integrations within categories).* This measure reflects the degree to which a person's strategies focus narrowly on defined groups of actions and plans. For example, all strategies that are specific and only relevant to caring for school children during a threatening disaster would be counted. However, if the disaster control coordinator concurrently planned to park these buses subsequently across streets to prevent cars from entering flooded areas, such an action sequence would not be counted. All integrations representing single intent sequences are summed.

8. *Systematic Functioning (total integrative activity).* This measure assesses the degree to which planning, strategic action, opportunistic action, and purposive repetitive actions are combined into an overall approach to

managerial effectiveness. Values for "Uses of Strategy," "Utilizing Opportunity," "General Planning," and a measure of "Purposive Repetitive Action" are added. The latter measure counts all action sequences (as described in the "Use of Strategy" measure) where both actions belong to a single category (as described in the "Diversity of Action" measure). Potential scores for "Systematic Functioning" are only restricted by the number of actions upon which a participant decides.

9. *Systematic Approach to Strategic Planning (total forward integrative activity).* This measures sums *only two* activities that are included in the previous measure: "Use of Strategy" and "General Planning." The measure assesses the degree of forward-directed strategic activity.

10. *Utilizing Strategy in Handling an Emergency (disaster total integrative activity).* The simulations introduce an emergency that requires immediate decisive action. The degree to which planning, strategic actions, utilization of opportunity and purposive repetitive actions (as discussed in the "Systematic Functioning" measure) are directly applied to a programmed emergency is evaluated. The potential score is restricted only by the number of decisions relevant to the emergency.

11. *Emphasis on Deliberate Strategic Thinking.* The degree to which multiple interrelated strategic plans are realized via rapid successive steps is considered. Excessive responsiveness to incoming information results in a reduction of the attained score. A value for this measure is obtained by diminishing the score for "High-Level Strategic Planning" on the basis of the average length of time across which actions are integrated and on the basis of the number of general (nonresponsive) decisions that are not part of any plan or strategy.

12 *Length of Forward Planning (time weight).* While the "Elapsed Time Between Strategic Actions" measure *sums* time periods between all integrated decisions (i.e., is sensitive to the "Use of Strategy" score), the present measure divides the "Elapsed Time" measure by the "Use of Strategy" score. As a result, an average value for elapsed time between strategically interrelated decisions is obtained.

Measures of Activity

13. *Response Frequency to Information (number of respondent decisions).* This score reflects dependence on incoming information to generate activity. The number of actions in direct response to incoming information is counted. Where several actions occur in response to single items of information, all of those actions are counted.

14. *Number of Actions in Response to an Emergency (number of disaster decisions).* Actions taken to deal with an emergency are summed. The

procedure is identical to the one discussed immediately above; however, only actions in response to emergency information are included.

15. *Simple Nonstrategic Actions (general unintegrated decisions).* The measure assesses the extent to which actions remain unrelated to plans and strategies (i.e., the extent to which they stand alone, potentially even out of context of task requirements). Any action that is not part of a plan or strategy is included (regardless of whether it is or is not responding to incoming information). The potential score may be partly restricted by the number of actions upon which a participant decides.

16. *Information-Search Activity (number of information-search decisions).* General search activity is not necessarily a good predictor of information utilization. However, information utilization is, to some extent, dependent on the frequency and effectiveness of search. The present measure reflects the number of times additional information was requested during the simulation. The potential score may be restricted by the number of actions upon which a participant decides.

Measures of Speed of Action

17. *Speed of Response to Information (average response speed).* Actions in direct response to information receipt are considered, including multiple responses to the same item of information. The average time between information receipt and action is calculated. Higher scores suggest that at least some responses occurred with some time delay.

18. *Rapidity of First Response to Information (most recent response speed).* The average elapsed time between arrival of information and the *first* response to that information is calculated.

19. *Speed of Response to Emergency Information (disaster response speed).* The measure calculates the average time between arrival of emergency information and actions in response to that information. A higher score represents a slower response to emergency action demands. The potential score may be modified somewhat by the frequency of responses to emergency information and by their specific timing.

20. *Shift in Speed upon an Emergency.* Under nonstressful conditions, one can often take more time to respond to incoming information. During an emergency such a delay can be disastrous. The present measure calculates changes in response rapidity as a disaster occurs. Disaster response speed is divided by speed of response to information.

TABLE 1
Reliability Values

Measure	r
1. Diversity of Action	.68
2. Use of Strategy	.80
3. Elapsed Time Between Strategic Actions	.74
4. Utilizing Opportunity	.80
5. High-Level Strategic Planning	.78
6. General Planning	.83
7. Strategic Actions within Groups of Coherent Activities	.69
8. Systematic Functioning	.94
9. Systematic Approach to Strategic Planning	.86
10. Utilizing Strategy in Handling an Emergency	.62
11. Emphasis on Deliberate Strategic Thinking	.81
12. Response Frequency to Information	.82
13. Number of Actions in Response to an Emergency	.43
14. Simple Nonstrategic Actions	.61
15. Information Search Activity	.14
16. Speed of Response to Information	.67
17. Rapidity of First Response to Information	.26
18. Shift in Speed Upon an Emergency	.11
19. Speed of Response to Emergency Information	.61
20. Length of Forward Planning	.74

Results

Measure intercorrelations for the two simulations are reported in Table 1. The intercorrelations across simulations for the 12 structural measures of managerial style ranged from $r = .62$ to $r = .94$. The majority of these intercorrelations ranged from .74 through .94. In other words, the obtained *alternate form* reliabilities were highly significant.

Measures of activity produced quite diverse intercorrelations, varying from $r = .82$ for response frequency to information to $r = .14$ for information-search activity. Measures of activity deal only to some extent with managerial style. They are also content sensitive (i.e., represent mixed measures). Partially stylistic actions are more easily modified by differences in response habits that might have originated during prior experience (e.g., in association with previous task content). The specific content characteristics of the two simulation scenarios likely had some impact on the obtained score and intercorrelations values.

Mixed measures of speed of action also produced variable r values ranging from .67 to near zero. Again, scenario content likely had some effect on the timing of responses. A split-half intercorrelation procedure based on some simpler scenario that would provide constant levels of load, success, failure, and so forth may provide more accurate information about reliability for mixed measures of managerial functioning.

Validity

Method

One hundred and eleven adults, again with managerial experience, participated in one or the other of the two simulations and, in addition, responded to a biographic interview. As part of the interview, a number of job success indices were obtained. They were (1) income at age (corrected for job category, gender, and location[2]), (2) job level at age, (3) number of persons supervised, and (d) number of promotions during the last 10 years. The latter four measures were employed as criterion variables for managerial success. Some caution in the interpretation of these indices of success is needed. The measures are based on self-reports that may generate self-enhancing bias by the respondents.

Managers in the sample were recruited from southcentral Pennsylvania. Average age was 43.13 (s = 10.0). They tended to occupy mid-level managerial positions. All had been employed as managers for at least 10 years, and none had risen to the level of corporate vice-president or higher. Most held masters or other post-baccalaureate degrees and listed their health as good with an average of .31 existing medical problems (s = .48). They had held their present jobs at an average of 4.62 (s = 3.22) years and worked for 1.81 (s = 1.80) organizations during the last 10 years. To obtain preliminary data on potential effects of satisfaction on performance, a single 7-point scale running from 1 = "extremely dissatisfied" to 7 = "extremely satisfied" was administered. Respondents rated satisfaction with their present job at an average of 4.50 (s = .89), their job level at 4.44 (s = 1.93), and their present salary level at 3.31 (s = 1.35).

To provide more pure indicators of the relationship between criterion variables and simulation performance, several measures were recalculated prior to validity analysis. Some of the 20 simulation measures (see above) contained components of other measures. For example, scores for "Use of Strategy" may potentially have been restricted by the number of participant actions. To avoid any introduction of bias that might have been contributed by validity coefficients of component measures, scores on several measures were divided by the value of the contributing measure. For

---

[2] Corrections were based on special tabulations of income at age purchased from the Bureau of Labor Statistics, U.S. Department of Labor. The Bureau produced the tables from their 1985 "Current Population Survey" microdata file showing estimated median annual earned income at age by location, gender, and occupation (based on the longest job held during the year). To obtain a corrected income value, the proportion of each respondent's (self-reported) earned income was divided by the median tabulated income at that person's age for the relevant occupation, job location and gender. The average nationwide annual income at the relevant age for the category "managerial occupations" was then multiplied with the obtained proportion.

example, participants' scores for "Use of Strategy" were divided by the "Number of Decisions Made." Where recalculations were employed, the resulting proportionate measure was identified by a (P) following its name.

Results

Intercorrelations of performance scores with four job success criteria and multiple regression beta weights were calculated. Correlational and beta weight data tended to be highly similar. Satisfaction scores tended to remain independent of success indices and were not related to simulation performance. Limited interrelationships between simulation scores for "Speed of Response to Information" and for "Length of Forward Planning" with educational level were obtained. Several simulation-based performance scores predicted success. Intercorrelations among simulation scores and criterion scores are provided in Table 2.

Intercorrelations among simulation measures show patterns similar to the major factors obtained in a previous analysis (Streufert et al., 1986). This paper will initially focus on two groups of measures that are concerned with the assessment of strategic functioning. The first, reflecting a previously extracted factor concerned with complex strategic actions, was centered around the "High-Level Strategic Planning" and "Use of Strategy" measures but included "Systematic Approach to Strategic Planning" as well. Further, the measure, "Elapsed Time Between Strategic Actions," was meaningfully associated with "Use of Strategy" but not with "High-Level Strategic Planning," the most complex indicator of strategic functioning.

The measure, "Systematic Approach to Strategic Planning," contains two components: "Use of Strategy" (a more complex action according to complexity theory) and the somewhat lower-order "General Planning" activity. (Note that plans counted by the latter measure might not be completed due to lack of opportunity/necessity, but also due to lack of competence.) It is then not surprising that the "Total Forward Integrative Activity" measure represents a bridge between measures reflecting higher versus those reflecting lower levels of strategic activity. The group of measures reflecting a previously extracted factor concerned with lower levels of strategic functioning also includes "Systematic Functioning," "Strategic Actions within Groups of Coherent Activities," "Utilizing Opportunity," "General Planning," and negatively, "Simple Nonstrategic Actions." Generally, somewhat lower correlations interrelate the two groups of measures. Both groups of strategy-oriented measures failed to intercorrelate meaningfully with educational level ($r < .20$). Both groups produced interrelationships with two criterion variables: occupational level and income at age.

TABLE 2

Intercorrelations Among Simulation Measures and Criterion Variables

A	B	C	D	E	F	G	H	I	J	K	L	
–												A. Diversity of Action (P)
21^a	–											B. Use of Strategy (P)
−04	51^b	–										C. Elapsed Time Between Strategic Actions
31^b	06	13	–									D. Utilizing Opportunity (P)
16	43^b	12	01	–								E. High−Level Strategic Planning (P)
30^b	28^b	02	−02	38^b	–							F. General Planning (P)
−06	31^b	32^b	39^b	38^b	24^a	–						G. Strategic Actions within Groups of Coherent Activities
−04	41^b	31^b	67^b	29^b	50^b	53^b	–					H. Systematic Functioning (P)
−16	54^b	38^b	38^b	40^b	44^b	55^b	69^b	–				I. Systematic Approach to Strategic Planning
12	−06	−16	25^b	11	36^b	19^a	29^b	13	–			J. Utilizing Strategy in Handling an Emergency (P)
18	21	12	−26^b	18	14	−03	−06	02	23^a	–		K. Emphasis on Deliberate Strategic Thinking
−45^b	−09	10	12	11	−20^a	13	03	13	−10	−01	–	L. Response Frequency to Information
−31^b	28^b	52^b	05	−06	−22^a	16	10	24^b	−31^b	−01	49^b	M. Number of Actions in Response to an Emergency
−05	−29^b	−21	−41^b	−25^b	−28^b	−38^b	−55^b	−31^b	−16	06	−09	N. Simple Nonstrategic Actions (P)
52^b	14	02	−05	04	18	11	10	−17	27^b	19^a	−08	O. Information-Search Activity (P)
16	07	03	−40^b	03	−07	27^b	27^b	−22^a	−08	53^b	25^b	P. Speed of Response to Information (average.)
01	03	01	−30^b	10	−04	−14	22^a	−10	14	56^b	20^a	Q. Rapidity of First Response to Information
24^b	−10	−15	−02	−07	00	04	−06	−20^a	−19^a	28^b	−35^b	R. Shift in Speed upon an Emergency
−06	10	38^b	05	−21^a	−23^a	08	01	−03	−44^b	24^a	03	S. Speed of Response to Emergency Information
04	−46^b	−30^b	−35^b	−26^b	−21^a	−41^b	−44^b	−46^b	−08	−04	02	T. Length of Forward Planning (P)
07	23^b	16	42^b	39^b	67^b	54^b	66^b	52^b	63^b	30^b	07	U. Occupational level at age
−04	−14	−20^a	13	24^a	−08	12	01	−19	01	−10	−30^b	V. Number of employees supervised
06	31^b	34^b	31^b	45^b	49^b	46^b	53^b	41^b	50^b	51^b	02	W. Income at age
41^b	−02	−43^b	08	−16	45^b	01	01	08	60^b	24^a	05	X. Number of promotions in 10 years

$^a p<.05$; $^b p<.01$

Table 2 (continued)

M	N	O	P	Q	R	S	T	U	V	W	X	
												A. Diversity of Action (P)
												B. Use of Strategy (P)
												C. Elapsed Time Between Strategic Actions
												D. Utilizing Opportunity (P)
												E. High–Level Strategic Planning (P)
												F. General Planning (P)
												G. Strategic Actions within Groups of Coherent Activities
												H. Systematic Functioning (P)
												I. Systematic Approach to Strategic Planning
												J. Utilizing Strategy in Handling an Emergency (P)
												K. Emphasis on Deliberate Strategic Thinking
												L. Response Frequency to Information
–												M. Number of Actions in Response to an Emergency
–12	–											N. Simple Nonstrategic Actions (P)
–10	–34^{b}	–										O. Information-Search Activity (P)
07	11	29^{b}	–									P. Speed of Response to Information (average)
–02	26^{b}	14	49^{b}	–								Q. Rapidity of First Response to Information
–44^{a}	–03	–11	–17	–12	–							R. Shift in Speed upon an Emergency
43^{b}	–11	–10	–06	23^{a}	28^{b}	–						S. Speed of Response to Emergency Information
–28^{b}	47^{b}	–05	16	24^{a}	10	–09	–					T. Length of Forward Planning (P)
–21^{a}	–58^{b}	24^{a}	–11	23^{a}	–12	–32^{b}	–35^{b}	–				U. Occupational level at age
37^{b}	–11	–15	–14	–20^{a}	41^{b}	–10	08	02	–			V. Number of employees supervised
–03	–47^{b}	26^{b}	05	33^{b}	–15	–14	–28^{b}	69^{b}	–03	–		W. Income at age
–50^{b}	–05	49^{b}	15	27^{b}	–16	–51^{b}	20^{a}	38^{b}	–25^{b}	31^{b}	–	X. Number of promotions in 10 years

$^{a}p<.05$; $^{b}p<.01$

A third group of intercorrelated measures relate to a previously obtained factor for response speed. It includes "Emphasis on Deliberate Strategic Thinking," "Speed of Response to Information," and "Rapidity of First Response to Information." Correlations with educational level varied from .07 to −.31. The first and third measures relate to the income at age, promotions, and occupational-level criteria.

A fourth group of three associated measures reflects a previously identified factor of disaster sensitivity, which primarily addresses whether participants responded quickly during an emergency. The three measures are "Speed of Response to Emergency Information," "Number of Actions in Response to an Emergency," and "Elapsed Time Between Strategic Actions." Correlations of two of these measures with educational level remained below $r = .2$. Speed of response to emergency information and educational level correlated at $r = -.40$. High scores on all three measures indicate slow responding. All three measures produced negative intercorrelations with number of promotions (from −.43 to −.51). Lesser interrelationships were obtained with other criterion variables.

A somewhat weaker interrelated group of measures had previously loaded on a decision activity factor. The measures were "Response Frequency to Information," "Number of Actions in Response to an Emergency," and "Shift in Speed upon an Emergency." Correlations between these measures and educational level remained below $r = .2$. "Response Frequency to Information" and "Number of Actions in Response to an Emergency" were negatively related to number of employees supervised ($r = -.30$ and $-.37$). In contrast, "Shift in Speed upon an Emergency" produced a positive relationship ($r = .41$) The "Number of Actions in Response to an Emergency" was negatively related to number of promotions (−.50).

The "Diversity of Action" measure, an indicator of differentiation, did not combine with other measures into any clearly identifiable group. However, it was correlated with "Information-Search Activity," with "Response Frequency to Information," and to a lesser extent with utilizing opportunity and "Number of Actions in Response to an Emergency." The latter measures were not meaningfully intercorrelated. "Information-Search," which, in turn, was also negatively correlated with "Simple Nonstrategic Actions," is often necessary to generate opportunistic or differentiated actions reflected in the diversity measure, while excessive responding to information consumes time that may hinder a diverse approach to problems. Both "Diversity of Action" ($r = .41$) and "Information-Search Activity" ($r - .49$) were related to the number-of-promotions criterion. Both measures correlated negatively with educational level ($r = -.29$ and $-.25$, respectively). While those relationships were not high, somewhat better performance by younger simulation participants may have been involved.

The measure, "Utilization of Opportunity," which was discussed in association with measures of strategy and diversity of action, generated additional relationships with "Speed of Response to Information" and with "Rapidity of First Response to Information." "Opportunistic Action" was apparently associated with quick responding and with a mixed group of criterion variables but not with educational level ($r < .20$).

"Utilizing Strategy in Handling an Emergency" was interrelated with two simulation measures that remained independent of each other ("Number of Actions in Response to an Emergency" and "Speed of Response to Information") as well as to three criterion variables (occupational level, $r = .63$; income at age, $r = .50$; and number of promotions in 10 years, $r = .60$). Intercorrelations with educational level remained below .20.

The "General Planning" measure interrelated generally with several simulation measures, especially those identified with strategic clusters and clusters focusing on speed and action frequency. Planning was also widely related to criterion variables ($r = .67$ with occupational level, .49 with income, and .45 with promotions) but not to educational level.

A contrast between the "Elapsed Time Between Strategic Actions" and the "Length of Forward Planning" measures may be of special interest. Both calculate the amount of time between any two strategically interrelated decisions. However, "Elapsed Time" sums those times (i.e., reflects in part, the degree to which a participant is involved in the use of strategy). "Length of Forward Planning" divides that elapsed-time value by the "Use of Strategy" score. As a result, effects of the "Use of Strategy" component are eliminated. "Length of Forward Planning" was negatively related to some strategy-oriented measures, to three of the criterion variables, and to educational level ($-.33$). Since "Elapsed Time" scores include the positive criterion relationships of the "Use of Strategy" measure and negative relationships found for "Length of Forward Planning," a mixed relationship to criterion variables emerged for that measure.

It should be noted that several of the simulation-based measures were associated with more than one criterion index. For example, "Rapidity of First Response to Information" was significantly interrelated with four success indices. "Elapsed Time Between Strategic Actions," "High-Level Strategic Planning," "General Planning," "Utilizing Strategy in Handling an Emergency," Emphasis on Deliberate Strategic Thinking," "Simple Nonstrategic Actions," "Information-Search Activity," and "Length of Forward Planning" were significantly correlated with three of the criterion measures. "Use of Strategy," "Utilizing Opportunity," "Strategic Actions within Groups of Coherent Activities," "Systematic Functioning," "Systematic Approach to Strategic Planning," "Simple Nonstrategic Actions," and "Shift in Speed upon an Emergency" related to two criteria. Three measures were significantly correlated with one criterion, and one ("Speed of Response to Information"/average) failed to relate to the criterion variables.

Discussion

The obtained data suggest that the two simulations generate reliable data. Reliability was demonstrated especially for measures of structural style, which focus, for example, on managerial planning and strategic action (i.e., on behaviors that are applicable to complex, uncertain, and fluid task environments). Mixed measures involving structural style *and* content orientation were, as one might expect, not as consistently reliable in this alternate form (multi-scenario) analysis. Special split-half procedures, involving especially designed scenarios, would have to be developed to test reliability for such measures.

The data suggest validity for several simulation-based measures. Several intercorrelations between simulation measures and criterion variables as high as $r = .50$ to .67 were observed. The correlations were, of course, based on the interrelationships between objective performance measures and subjective (self-report) indicants of success. Self-reports may be subject to some self-enhancing bias. To confirm the obtained external validity, future research might interrelate simulation performance with measures of success that are externally provided. Longitudinal research designs that permit inference of causality and the replacement of self-report criterion measures with outside criteria would be especially useful.

One interesting finding of this research should be emphasized. The measures "Length of Forward Planning" and "Simple Nonstrategic Actions" were negatively related to several performance scores and to job success. Such a negative relationship would certainly be expected for "Simple Nonstrategic Actions." This measure reflects performance that would not aid strategic or planning functions and may even be unrelated to concurrent task demands. However, the finding that "Length of Forward Planning" is *negatively* associated with success would potentially disagree with the views of Jaques and associates (e.g., Jacques, 1976), who have argued that length of planning should be an excellent indicator of managerial competence. The present data would argue otherwise (not even considering the fact that length of planning produced a moderately negative correlation with educational level). It is possible that successful individuals do indeed take shorter-term views of their task environment. Short-term views would, however, hardly prevent the development and execution of intentional *sequences* of short-term strategies. Likely, a pure length-of-planning concept may have to be replaced with sequential steps of strategic interrelated planning, which (cf. the measure of "High-Level Strategic Planning") could, in at least some executive jobs, permit planning sequences that might extend

MEASURE	Correlation of r > .3 With Educational Level	Alternate Form Simulation Reliability of r > .60	PREDICTIVE VALIDITY: Number of Promotions/10 Years r > .30	Income at Age r > .30	Occupation Level at Age r > .30	Number of Employees Supervised r > .30	Significantly Improved Performance With Training	Significantly Diminished Performance With Age	Significant Effects of Prescription Drugs on Performance
Diversity of Action (P)		■	■					■	■
Use of Strategy (P)		■		■			■	■	■
Elapsed Time Between Strategic Actions		■	▨	■					
Utilizing Opportunity (P)		■		■	■				■
High Level Strategic Planning (P)		■		■	■		■	■	■
General Planning (P)		■	■	■	■		■		
Strategic Actions Within Groups of Coherent Activities		■		■	■			■	
Systematic Functioning (P)		■		■	■		■	■	
Systematic Approach to Strategic Planning		■		■	■		■	■	■
Utilizing Strategy in Handling an Emergency		■	■	■	■		■	■	
Emphasis on Deliberate Strategic Thinking		■		■	■				
Response Frequency to Information		■				▨			
Number of Actions in Response to an Emergency			▨			▨			
Simple Non-Strategic Actions (P)		■		▨	▨			■	
Information Search Activity (P)			■					■	
Speed of Response to Information (Average)		■							
Rapidity of First Response to Information				■			■		
Shift in Speed Upon an Emergency						■	■		
Speed of Response to Emergency Information	▨	■	▨		▨		■		
Length of Forward Planning (P)	▨	■			▨			■	

Figure 1: **Association Among Simulation Measures and Other Selected Variables** (Blacked-out squares indicate positive asssociations between variables. Shaded squares indicate negative associations.)

for considerable periods of time. Such a modification would not violate the spirit of Jaques' theory.

With greater assurance of reliability and validity, and with measurement extended considerably beyond prior limitations imposed by a more restricted focus of complexity theory, it may now be possible to employ quasi-experimental simulation technology for assessment, training, and other purposes. Some success-of-training efforts based on the simulation technology have already been demonstrated (cf. Streufert, Pogash, & Piasecki, 1986). Other concurrent research has considered the effects of managers' age on performance and the impact of beta-blocking drugs on managerial performance (Streufert, Pogash, Piasecki, & Repman, 1986). A summary, which contains an overview of the results reported in the present paper as well as some results from other concurrent research, is provided as Figure 1.

REFERENCES

Bhutani KA. (1977). A study of the effect of some cognitive and personality factors on attitude change. *Indian Educational Review*, *12*, 50–60.

Byham WC, Thornton GC III. (1986). Assessment centers. In Berk RA (Ed.), *Performance assessment: Method and applications* (pp. 143–166). Baltimore: Johns Hopkins University Press.

Fromkin HL, Streufert S. (1976). Laboratory experimentation. In Dunnette M (Ed.), *Handbook of organizational and industrial psychology* (pp. 415–465). Chicago: Rand McNally.

Hale CL. (1980). Cognitive complexity-simplicity as a determinant of communication effectiveness. *Communication Monographs*, *47*, 304–311.

Hendrick HW. (1979). Differences in group problem solving behavior and effectiveness as a function of abstractness. *Journal of Applied Psychology*, *64*, 518–525.

Heriot P, Chalmers C, Wingrove J. (1985). Group decision making in an assessment centre. *Journal of Occupational Psychology*, *58*, 309–312.

Jaques E. (1976). *A general theory of bureaucracy*. London: Heinemann Educational Books.

Karlins M, Lamm H. (1967). Information search as a function of conceptual structure in a complex problem solving task. *Journal of Personality and Social Psychology*, *5*, 456–459.

Linville PW, Jones EE. (1980). Polarized appraisals of out-group members. *Journal of Personality and Social Psychology*, *38*, 689–703.

Mizerski RW. (1978). Causal complexity: A measure of consumer causal attribution. *Journal of Marketing Research*, *15*, 220–228.

Russell CJ. (1985). Individual decision processes in an assessment center. *Journal of Applied Psychology*, *70*, 737–746.

Schroder HM. (1982). *Conceptual complexity and managerial performance*. Presented to the International Congress of Applied Psychology, Edinburgh, Scotland.

Scott WA, Osgood DW, Peterson C. (1979). *Cognitive structure: Theory and measurement of individual differences*. Washington, DC, and New York: V. H. Winston & Sons and Halsted Division, John Wiley & Sons.

Streufert S. (1966). Conceptual structure, communicator importance and interpersonal attitudes toward deviant and conforming group members. *Journal of Personality and Social Psychology*, *4*, 100–103.

Streufert S. (1970). Complexity and complex decision making: Convergences between differentiation and integration approaches to the prediction of task performance. *Journal of Experimental Social Psychology*, *6*, 494–509.

Streufert S. (1978). The human component in the decision making situation. In King BT, Streufert S, Fiedler F (Eds.), *Managerial control and organizational democracy* (pp. 215–230). Washington, DC, & New York: V. H. Winston & Sons and John Wiley & Sons.

Streufert S. (1986). How top managers think and decide. *Executive Excellence, 8,* 7–9.

Streufert S, Driver MJ. (1965). Conceptual structure, information load and perceptual complexity. *Psychonomic Science, 3,* 249–250.

Streufert S, Driver MJ. (1967). Impression formation as a measure of the complexity of conceptual structure. *Educational and Psychological Measurement, 27,* 1025–1039.

Streufert S, Pogash RM, Piasecki MT. (1986). *Training for cognitive complexity.* Final report to the U.S. Army Research Institute for the Behavioral and Social Sciences.

Streufert S, Pogash RM, Piasecki MT, Repman MA. (1986). *Age and human performance: A preliminary report on data.* Report to the U.S. Government.

Streufert S, Pogash RM, Piasecki MT, Repman MA, Swezey RW. (1986). *Data collection via a quasi-experimental simulation, III: Factor structure and validity.* Report to the U.S. Army Research Institute for the Behavioral and Social Sciences.

Streufert S, Streufert SC. (1969). Effects of conceptual structure, failure and success on attributions of causality and interpersonal attitudes. *Journal of Personality and Social Psychology, 11,* 138–147.

Streufert S, Streufert SC. (1978). *Behavior in the complex environment.* Washington, DC, and New York: V. H. Winston & Sons and Halsted Division, John Wiley & Sons.

Streufert S, Suedfeld P, Driver MJ. (1965). Conceptual structure, information search and information utilization. *Journal of Personality and Social Psychology, 2,* 736–740.

Streufert S, Swezey RW. (1985). Simulation and related research methods in environmental psychology. In Singer J, Baum A (Eds.), *Advances in environmental psychology* (vol. 5, pp. 99–117).

Streufert S, Swezey RW. (1986). *Complexity, managers, and organizations.* New York: Academic Press.

Suedfeld P. (1964a). Attitude manipulation in restricted environments, I. Conceptual structure and response to propaganda. *Journal of Abnormal and Social Psychology, 68,* 242–247.

Suedfeld P. (1964b). Conceptual structure and subjective stress in sensory deprivation. *Perceptual and Motor Skills, 19,* 896–898.

Suedfeld P. (1978). Characteristics of decision making as a function of the environment. In King B, Streufert S, Fiedler F (Eds.), *Managerial control and organizational democracy* (pp. 203–213). Washington, DC, & New York: V. H. Winston & Sons and Halsted Division, John Wiley & Sons.

Suedfeld P, Corteen RS, McCormick C. (1986). The role of integrative complexity in military leadership: Robert E. Lee and his opponents. *Journal of Applied Social Psychology* (special issue on military psychology), *16,* 498–507.

Suedfeld P, Rank AD. (1976). Revolutionary leaders: Long-term success as a function of changes in conceptual complexity. *Journal of Personality and Social Psychology, 34,* 169–178.

Suedfeld P, Tetlock PE. (1977). Integrative complexity of communications in international crisis. *Journal of Conflict Resolution, 21,* 169–184.