The conventional paradigm for testing expectancy theory predictions of work behavior has been to correlate expectancy-value reports with concurrent measures of motivation and performance. Although this static, two-variable approach has typically yielded statistically significant results, correlations have not been sizable. This study, using a sample of design and development engineers over a four-year period, posited and found that expectancy theory predictions were materially strengthened by taking into account the effects of five complicating factors. Specifically, these factors were: (1) time, (2) the initial level of the criterion, (3) the level of rewards, (4) task-specific ability, and (5) control system responsiveness. (Author)
FACTORS COMPLICATING EXPECTANCY THEORY PREDICTIONS
OF WORK MOTIVATION AND JOB PERFORMANCE

Richard E. Kopelman
Baruch College
The City University of New York

The conventional approach to testing expectancy theory has been to correlate expectancy-value reports, on the one hand, with contemporaneous measures of individual work motivation and job performance, on the other. This paper focuses on five complicating factors which, when taken into account, materially improved the predictive power of expectancy theory. Specifically, these factors are: 1) time, 2) initial level of criterion, 3) level of rewards received, 4) control system responsiveness, and 5) task-specific ability. These factors are considered in turn, after a brief description of the research methodology employed.

METHOD

The sample in 1969 consisted of 399 design and development engineers employed by three companies (Company 1, n=142; Company 2, n=138; Company 3, n=119). Four years later, 258 of the original 399 were still employed by the same three companies, and of these 210 responded to the 1973 questionnaire (Company 1, n=59; Company 2, n=56; Company 3, n=95). The attrition experience in the sample from each company was virtually identical with that of the population from each company.
Individual job performance was measured directly by supervisory performance rankings which were generated by ongoing organizational practices. The one exception: Company 2 had no formal evaluation procedure in 1969, so engineering managers were asked to rank engineers for this research. By 1971, however, all three companies utilized a formal evaluation procedure. Two less direct measures of performance included salary and organizational level (job title).

Motivation was measured by self reports of effort and hours worked in 1969 and 1973.

Expectancies were operationalized by the question, "What are your chances?" Responses were obtained using a 7-point scale with the end scores labeled "excellent" and "poor". This instrument measured overall expectancies, rather than the "artificially constricted" performance-based dimensions [Mitchell, p. 32, in press] of E1 (effort-performance) and E2 (performance-reward) components. Expectancies (and values) were obtained for five outcomes: three extrinsic outcomes "promotion to the next engineering grade", "promotion into management", "significant salary increases based on merit", and two intrinsic outcomes "more challenging work assignments", and "more freedom to carry out your own ideas and to use your originality and initiative".

Value scores were obtained for each outcome using the question, "How important is this to you?" Again, a 7-point scale was used, with the end values labeled "important" and "unimportant". Expectancy and value questions were alternated for each outcome.
RESULTS AND DISCUSSION

Time

Numerous writers (e.g., Porter and Lawler, 1968; Mitchell and Albright, 1972) have pointed out that expectancy theory is a temporal model. It takes time for a person's work preferences as operationalized by expectancy-value reports (hereafter labeled EEV)\(^2\) to be translated into corresponding levels of motivated behavior, job performance, evaluated performance, and ultimately, rewarded performance. Yet, despite wide agreement that expectancy theory is not atemporal or syndromatic -- that is, all variables do not change instantaneously -- all prior expectancy studies, with four exceptions, have examined contemporaneous (static) correlations. The four previous longitudinal studies, Lawler (1968), Lawler and Suttle (1973), Sheridan, Slocum and Richards (1974), and Van Maanen (1972) examined correlations between EEV and subsequent performance over intervals of one year, one year, fifteen months, and seven months, respectively. The present research examined expectancy theory predictions over intervals of one, two, three, and four years, in addition to obtaining static correlations at two points in time.

It was found that the correlation between EEV and supervisory performance rankings was strongest after an interval of one year (\(r = .24\) static versus \(r = .45\) one-year later). In comparison, predictions of salary and organizational level were found to be strongest after a four-year interval -- perhaps explaining why no prior expectancy research has used either of these two criteria as dependent variables. Also,
given the fact that in the three companies studied performance rankings served as a basis for promotions and raises, the longer lag between IEV and these latter criteria is not surprising.

Lastly, utilizing a cross-lagged panel analysis, support was found for the notion that the direction of primary determining influence was from IEV to performance, not vice versa. After controlling for initial and subsequent levels of the criterion -- by examining results for "criterion cohorts" -- 89 out of 100 comparisons of static and longitudinal correlations were as predicted.

**Initial Level of Criterion**

Because this research obtained longitudinal performance data, it was possible to control for the initial level of all three criterion variables: performance rankings, salary, and organizational level. Several reasons were advanced why it might make sense to do this. First, with respect to all three criteria, there are floors and ceilings. Consequently the use of longitudinal data introduces the problem of regression to the mean. Secondly, existing ceilings and floors are well known to respondents, particularly with respect to salary and organizational level, and this knowledge may affect expectancy predictions. People with high salaries and organizational levels may have low expectancies for goal attainment simply because there is little room for advancement. On the other hand, people with low salaries and organizational levels may have high expectancies since the "only way to go is up." Third, controlling for initial cri-
Criterion levels minimizes the problem of control system bias or inertia. That is, it may take time for a person’s performance rankings, salary, or organizational level (particularly the latter two) to "catch up" so to speak with actual performance. For example, it is not uncommon for high performing younger engineers to have lower salaries and organizational levels than low performing older engineers.

Two methods were used to control for initial criterion levels. First, the population was segmented into subgroups comprised of "criterion cohorts", and expectancy predictions were analyzed for each subgroup. Second, first-order correlations were computed between $E_E$V and subsequent measures of each criterion, partialling out the initial level of the criterion. Both approaches yielded identical results, which were: predictions of performance rankings were basically unaltered by controlling for initial levels, but predictions of salary and organizational level were greatly strengthened.

$E_E$V was most predictive of performance rankings one year later ($r=.45$), but when the initial level was partialled out, the correlation was $r=.43$. In contrast, $E_E$V was most predictive of salary and organizational level after a four-year interval ($r=.06$ and $r=.10$, respectively). However, after initial levels of salary and organizational level were partialled out, the two correlations rose to $r=.53$ and $r=.27$, respectively.

As indicated above, the improvement in predictions of salary and organizational level resulting from partialling out initial criterion levels is consistent with 1) the existence and knowledge of ceilings and floors, and 2) control system lags in rewarding performance.
Level of Rewards

This research refrained from making any specific predictions regarding the comparative strength of performance predictions using $\$EV$ for extrinsic rewards versus $\$EV$ for intrinsic rewards. The rationale for this was that expectancy theory, unlike the various need theories, does not take a position as to which broad types of needs most motivate work behavior. Moreover, previous research has been equivocal with respect to the relative strength of extrinsic and intrinsic predictors.

Across all respondents it was found that extrinsic $\$EV$ was a better predictor than intrinsic $\$EV$: in 1969 $r = .24$ and $r = .16$ for extrinsic and intrinsic $\$EV$, respectively, and in 1973 the two correlations were $r = .40$ and $r = .11$, respectively. However, when the data were analyzed controlling for the level of rewards received (specifically, controlling for pay), some interesting results were obtained. For respondents whose salaries were among the lowest one-third, extrinsic $\$EV$ was a far better predictor of performance than was intrinsic $\$EV$: $r = .26$ versus $r = -.04$ in 1969; $r = .47$ versus $r = .03$ for 1970 performance (longitudinal); and $r = .59$ versus $r = .05$ in 1973. Among respondents whose salaries were among the highest one-third, though, extrinsic $\$EV$ was only a slightly better predictor than intrinsic $\$EV$ in two cases out of three: $r = .29$ versus $r = .26$ in 1969; $r = .45$ versus $r = .43$ for 1970 performance (longitudinal) and $r = .37$ versus $r = .09$ in 1973. Thus, after controlling for the level of rewards received, a Maslovian interpretation was suggested: extrinsic rewards were most predictive for respondents with a low level of extrinsic need satisfaction (inferred), and intrinsic rewards were most predictive for respondents whose extrinsic needs were presumably better satisfied.
Lastly, prior studies comparing extrinsic and intrinsic \( \text{LEV} \) predictions were reinterpreted along Maslovian lines. It may not have been entirely coincidental, for instance, that intrinsic \( \text{LEV} \) yielded higher correlations than extrinsic \( \text{LEV} \) among naval aviation officers (Mitchell and Albright, 1972), whereas the opposite result was found for machine operators during the late 1950's (Georgopoulos et al., 1957).

In short, it appears that controlling for levels of rewards (pay) may represent a difference which makes a difference. But in contrast to the position of House and Wahba (1972), that intrinsic rewards are more predictive of performance than extrinsic rewards, this paper advances a contingency approach. Whether extrinsic or intrinsic rewards are better predictors of performance may depend on the level of rewards, or more precisely the level of extrinsic need satisfaction, respondents have attained.

**Control System Responsiveness**

It was hypothesized that control system responsiveness -- the degree to which rewards (e.g., salary, organizational level) are contingent upon performance -- would be positively related to the strength of expectancy predictions of performance. The rationale for this premise was two-fold: control systems which were relatively more responsive to performance would 1) minimize the effects of non-performance-related factors on reward attainment, and 2) would strengthen the actual and perceived connection between motivated behavior and organizationally mediated rewards. Therefore, in an organization with high control system
responsiveness (CSR), expectancy-value reports (€EV) should be more highly correlated with performance than in an organization with low CSR.

As predicted, CSR was found to positively affect expectancy theory predictions of performance. In seven cases out of eight the relative magnitude of performance predictions corresponded to the relative level of CSR. Moreover, some of the differences in correlations were both statistically significant and sizeable: for example, in 1973 predictions were $r = .50$ (p < .01), $r = .21$ (n.s.), and $r = .16$ (n.s.) for Companies 3, 2, and 1, respectively.

At least five previous expectancy studies have examined the moderating effects of situational performance-reward contingencies (a concept analogous to CSR) on expectancy predictions of performance. With one exception, a positive relationship was found. A unique feature of the present study, though, is that control system responsiveness was measured both objectively (by computing static and longitudinal performance-reward correlations) and subjectively (by utilizing a factor analyzed questionnaire).

It was also posited that organizational differences in CSR would affect organization-wide levels of expectancy-value reports (€EV). Again the results were as predicted: in six out of six cases, differences across companies in CSR paralleled differences in €EV. Furthermore, differences in €EV between the two companies with the highest and lowest CSR scores were highly significant ($t = 2.84$, $p < .002$ in 1969 and $t = 5.08$, $p < .001$ in 1973). It might be noted that no prior non-experimental research has
examined the relationship between CSR and the level of EV. Yet the practical implications of this relationship are of considerable importance.

**Task-Specific Ability**

Whether or not a person can translate his behavioral preferences (EV) into actual job performance depends on his ability, the difficulty of the task, and most importantly on the intersection of these two factors. Based on Heider's (1958) concept of "can", which refers to the meshing of internal and external forces, moderator effects might be posited for ability and task difficulty, as well as the interaction between the two — task-specific ability. More specifically, a person's power or "can" should be positively related to his ability and negatively related to the difficulty of the task. Thus a 2x2 matrix was proposed in which motivational predictions of performance were posited to be a) strongest when both factors were favorable (high ability—low task difficulty); b) moderate when one factor was favorable (high ability—high task difficulty or low ability—low task difficulty), and c) weakest when neither factor was favorable (low ability—high task difficulty). The results of both separate moderators and their interaction are reviewed in turn.

**Task difficulty**

As hypothesized, task difficulty exhibited a moderator effect on motivational predictions of job performance. Dividing the sample into three subgroups of equal size, expectancy predictions were r=.07 in the high task difficulty subgroup, r=.14 in the moderate task difficulty subgroup, and r=.27 in the low task difficulty subgroup.
This finding is consistent with the one previous study (Lifter, Bass and Nussbaum, 1971) which examined expectancy predictions across tasks of differing difficulty. Arguing that the task of being a line supervisor was easier and more certain than the task of being a staff supervisor, Lifter et al reported motivational predictions of \( r = .10 \) for line supervisors and \( r = -.07 \) for staff supervisors.

Additional support for the moderator effect of task difficulty comes from examining performance predictions across expectancy studies. Although it is hazardous to compare results across studies, because other factors vary besides tasks, the pattern of results is remarkably consistent. Of twenty-four studies surveyed, the two reporting the highest (concurrent) product-moment correlations involved relatively simple and repetitive tasks -- machine operation (Schwab and Dyer, 1973) and a number processing exercise (Craen, 1969) -- whereas the weakest predictions involved relatively complex and variable tasks -- audit and tax accounting (Todd, 1972), college G.P.A. (Mitchell and Nebeker, 1971), staff supervision (Lifter et al, 1971). In short, there appears to be consistent support for task difficulty as a moderator of performance predictions.

**Individual ability**

In this research, ability was determined indirectly. The average (mean) performance ranking of engineers was computed for five-year age intervals, and within each interval engineers whose performance rankings were above the "age cohort" mean were viewed as having high ability; those ranked below the mean were placed in the low ability.
category. Expectancy predictions of performance among the high ability and low ability subsamples were $r = .32$ and $r = .23$, respectively. This finding is consistent with the results of Lawler (1966), the only prior study to use ability as a moderator of expectancy predictions of job performance. Lawler (1966) obtained correlations of $r = .30$ and $r = .11$ among high and low ability government supervisors, respectively.

**Ability and task difficulty**

As Heider noted, ability *per se* is only a partial determinant of the outcome of behavior — rather, performance depends on the meshing of internal and external forces. In other words, a person with superior ability might perform poorly on a very difficult task, whereas a person with inferior ability might "sparkle" at an easy task.

Using the aforementioned 2x2 design to analyze the moderating effects of task-specific ability on (concurrent) expectancy predictions, the results were as hypothesized. When ability was determined by performance ranking halves, expectancy predictions of performance were $r = .42$ in the most favorable condition (high ability-low task difficulty), and $r = -.08$ in the least favorable condition (low ability-high task difficulty). And as hypothesized, predictions in the two intermediately favorable conditions were in-between: $r = .22$ and $r = .26$.

The fundamental point is that the accuracy of motivational predictions of performance depends on the extent to which individuals can translate their efforts into job results. This translation process
depends on both the person's ability and the difficulty of his task assignment. Thus, it may be no coincidence that expectancy predictions have consistently been strongest for simple, repetitive tasks and weakest for complex, uncertain tasks. Conversely, it may not be a coincidence that in the one non-experimental study in which ability scores were more predictive of performance than was \( E \) (Mitchell and Nebeker, 1971), the task was particularly difficult, involving college grade point average.

CONCLUSION

It has been found that expectancy predictions were significantly affected by a number of complicating factors; and hence, the relationship between independent and dependent variables should not be examined in vacuo. As Lawler (1971) has written, "Perhaps the clearest implication of the model [expectancy theory] is that... future research must be more complex and must look at more than two variables. Two-variable research simply is not appropriate, given the complexity of the relationships among the variables... Multiple variable research and thinking are admittedly more difficult, but they also are potentially more rewarding because of the increased explanatory power they offer." [p. 272]

In short, notwithstanding the proliferation of static, two-variable expectancy studies and the subsequent disappointment of some writers, the predictive validity of expectancy theory ultimately depends on the manner in which the theory is tested. After all, an unopened umbrella can hardly fail to be a disappointment to its user.
1. This research was supported by grants from the Graduate School of Business, Harvard University, and the William B. Harding Foundation.

2. In this paper the mathematical expression (iEV) refers to a generalized expectancy-value predictor. This notation is not meant to describe a particular operational definition of an expectancy-value predictor, even when used with regard to a specific study. Rather, it simply denotes the traditional summing of the products of expectancies and values, however they are operationalized— including VIE models, expectancy x valence formulations, and so forth.

3. Using the FCP (frequency of change in product-moment correlations) technique to assess causal priorities, similar results were obtained. Specifically, \( \chi^2 = 14.24 \) that the direction of primary determining influence was from iEV to performance, not vice versa.

4. These data are for respondents from Companies 2 and 3 only. Results were not supportive when analyzed across all respondents.

5. These data are for respondents from Companies 2 and 3 only. Results were supportive, but to a lesser degree, when analyzed for all respondents.
REFERENCES


