Journal of Applied Psychology
1985, Vol. 70, No. 4, 651-661

0021-9010/85/$00.75

An Examination of Two Alternative Techniques to Estimate the Standard Deviation of Job Performance in Dollars

Richard R. Reilly and James W. Smither
Applied Psychology Program
Stevens Institute of Technology

Two methods for estimating dollar standard deviations were investigated in a simulated environment. Subjects, with management experience, managed a simulated pharmaceutical firm for 4 quarters. Over a period of 4 weeks, subjects were given information describing the performance of sales representatives on three job components. The information was constructed by the experimenters to be normally distributed and was of known value. Methods proposed by Schmidt et al. and Cascio were used in counterbalanced order to obtain estimates of the dollar standard deviation of overall job performance. In addition, the Schmidt et al. procedure was used to estimate the dollar standard deviation of three job components. It was found that Schmidt et al. estimates were relatively accurate with objective sales data that could be directly translated to dollars, but resulted in overestimates of means and standard deviations when data were less directly translatable to dollars, and involved variable costs. An additional problem with the Schmidt et al. procedure involved the presence of outliers, possibly caused by differing interpretations of instructions. The Cascio–Ramos estimate of performance in dollars (CREPID) technique, proposed by Cascio, yielded smaller dollar standard deviations, but results showed that raters could reliably discriminate among job components in terms of importance and could accurately evaluate employee performance on those components. Problems with the CREPID method included the underlying scale used to obtain performance ratings and a dependency on job component intercorrelations.

Recently, there has been considerable interest in identifying the organizational benefits derived from the use of valid personnel selection procedures and other human resource interventions. Utility models proposed years ago by Brogden (1949) and Cronbach and Gleser (1965) have languished for a lack of acceptable methodology for estimating the monetary worth of employee performance. However, new techniques for estimating the standard deviation of job performance in dollar terms (SDy) have suggested the possibility that utility models can be applied to a wide variety of human resource problems. Three specific techniques for estimating SDy are those proposed by Schmidt, Hunter, McKenzie, and Muldrow (1979), Hunter and Schmidt (1983), and Cascio (1982).

The Schmidt et al. procedure is based on the assumptions that job performance in dollars is normally distributed and that experts (e.g., supervisors) can estimate the difference between the value of products and services produced by the average employee and those produced by an employee at the 85th (or 15th) percentile. The difference between the average value and the value of a performer at the 85th (or 15th) percentile is then used to estimate SDy.

A second method is based on the empirical findings (Hunter & Schmidt, 1983) that labor costs represent roughly half the cost of output and that the standard deviation of output is approximately 20% of mean output or 40% of average annual salary. Although the simplicity of this technique is appealing, Bobko, Karren, and Parkington (1983) argued that the Hunter and Schmidt method theoretically contradicts the normality assumption because it creates covariation between estimates of

We wish to thank Fred W. Cleveland for help in designing and conducting this study and the reviewers for their constructive criticisms of an earlier draft of the article.

Requests for reprints should be sent to Richard R. Reilly, Department of Management, Stevens Institute of Technology, Castle Point, Hoboken, New Jersey 07030.

mean productivity and estimates of the standard deviation of productivity.

The Cascio–Ramos estimate of performance in dollars (CREPID; Cascio, 1982) uses traditional principles of job analysis and performance measurement and is based on the economic premise that the value of an individual's labor is equivalent to what an organization is willing to pay to obtain that labor. Cascio and Ramos (1984) noted that they are not trying to equate wages with the gross dollar value of productivity or with the selling price of goods and services produced. Rather, they use wages as a baseline from which to measure the relative value of individual job performance to the firm. They note that an additional advantage of their approach is that, at worst, CREPID underestimates the value of labor.

The CREPID procedure involves eight steps. First, principal job activities are identified. Second, each principal activity is rated in terms of time/frequency, importance, consequence of error, and level of difficulty. Third, the numerical ratings for time/frequency, importance, consequence of error, and level of difficulty are multiplied for each principal activity. Fourth, dollar values are assigned to each principal activity by taking the average rate of pay of employees and allocating it across the principal activities according to the results of Step 3. Fifth, each employee is rated on each principal activity using a 0–200 point scale. Sixth, the dollar value of each principal activity is multiplied by the employee's point rating on that activity (expressed as a decimal number). Seventh, the overall economic value of each employee's job performance is computed by adding together the results of Step 6 for each principal activity. Eighth, over all employees in the study, the mean and standard deviation of dollar-valued job performance is computed. The CREPID procedure requires only two sets of ratings from supervisors; the job analysis ratings (Step 2) and the performance appraisal ratings (Step 5). The remaining steps can be completed by personnel specialists.

Several studies, using one or more of these procedures to estimate SDy, have demonstrated substantial utility for selection and training interventions for various jobs, including computer programmers (Schmidt et al., 1979), first-level telecommunication managers (Cascio & Ramos, 1984), and convenience store managers (Weekley, Frank, O'Connor & Peters, 1983).

Although these study results provide impressive evidence of the utility of human resource interventions, several issues concerning the SDy estimation procedures and the associated utility estimates deserve attention. Boudreau (1983) noted that utility formulas have failed to account for the effects of variable costs, taxes, and discounting on the utility of human resource intervention strategies; for example, increases in sales revenue or productivity may also increase variable costs. Sales representatives selected via an improved selection procedure may indeed increase sales revenue. However, these sales representatives may receive higher pay in the form of a commission or bonus. In addition, increased sales levels may require larger inventories or increased material costs and production costs. Boudreau (1983) noted that, in utility analysis, these increases in variable costs should be subtracted from the increased sales revenue resulting from increased productivity. Boudreau (1983) concluded that previous utility formulas are deficient and can produce upwardly biased utility estimates.

A more fundamental concern is the accuracy or validity of the SDy estimates obtained from the CREPID and Schmidt et al. (1979) procedures. If estimates of SDy are inaccurate, the effects on utility will be directly proportional.

In one empirical study of the accuracy of SDy, Bobko et al. (1983) utilized the Schmidt et al. (1979) procedure to obtain supervisors' estimates of the standard deviation of yearly dollar sales of insurance premiums. They compared these estimates to archival data and found that the SDy estimates were quite close to the actual standard deviation. Bobko et al. (1983) also noted that the supervisors underestimated actual percentile values. The estimated 50th percentile of $96,000 was significantly less than the actual median of sales data ($117,300) and the mean of sales data ($124,882). Although the point estimates were inaccurate, the effect of underestimation was removed when the difference between percentiles was used to determine SDy. As a result, the SDy estimates were quite accurate.

The findings of Bobko et al. (1983) suggested that subjects underestimate actual percentile values but still yield accurate SDy estimates of highly familiar, objective sales data. The ability of subjects to accurately estimate percentile values (or subsequent SDy estimates) of less familiar or explicit distributions, such as overall value of products and services, is unknown.

The purposes of this study were (a) to investigate the accuracy of SDy estimates using the Schmidt et al. procedure, (b) to examine the reliability and validity of the intermediate steps in the CREPID method, specifically, the assignment of importance to components and the rating of performance on those components, and (c) to demonstrate the differences that should result in using the two proposed methods for estimating SDy. The Weekley et al. (1983) study is illustrative of the dramatic impact that method related differences can have on estimated utility.

Subjects were provided with precise information concerning the job performance of 10 sales representatives on three primary job components (selling established product line, new product sales activity, and controlling expenses). For two of the job components, performance information was explicitly provided in terms of the dollar value of sales or expenses. For the third component (new product sales activity), performance information was provided in terms of the number of sales contacts and the percentage of sales contacts that were successful. Subjects were also provided with a simple formula used to translate the sales activity data on this third dimension into a dollar value of sales. In addition, subjects were provided with information concerning salary, commission, and variable costs (production expenses and overhead burden). Thus subjects were provided with precise employee performance information and operating expenses—information that would be available to a supervisor only under relatively ideal circumstances. For the CREPID method, the validity of rated performance could be assessed by correlations with actual performance. Reliability of the rated importance of components could also be checked. Using the Schmidt et al. (1979) procedure, estimates of SDy were obtained from subjects for the dollar value of sales of established product line (i.e., repeat sales), the dollar value of new product sales, the net sales revenue after expenses, and the value of the overall products and services produced (i.e., overall worth). Because the exact value of SDy was known for the first three of these estimates, the relative accuracy of subjects' estimates of several performance distributions could be assessed.

It should be noted that the Schmidt et al. and the CREPID method could not be directly compared, except on the overall SDy estimates. Schmidt et al. estimates of SDy for repeat sales, new product sales, and net revenue after expenses were attempts to directly estimate the SDy for these performance variables. In contrast, the CREPID estimate of SDy for a component of job performance, such as repeat sales, is assumed to be made in terms of that component's contribution to overall worth, which is, in turn, directly linked to average salary.

Method

Subjects

Study participants were enrolled in a graduate management class at an engineering and technical college. The class length was 14 weeks. All students had at least 2 years of management experience. The average length of supervisory experience for the subjects was 2.9 years. Subjects included 16 men and 3 women.

The course was based on a detailed interactive computer simulation, involving the management of a pharmaceutical company. The 19 subjects were divided into four groups; each group served as a separate team of executive managers provided with extensive financial (budgets, assets, liabilities, income statements, etc.) and product information (sales, production cost, inventory, equipment investment, new product possibilities, etc.). Each week, based on this information, each group made decisions required to operate the firm for several fiscal quarters. Typical decisions concerned capital structure, dividend policy, budgets (research, construction, promotion, distribution, and expansion), as well as construction outlay and production schedules for new and existing products. Students also dealt with general economic, legislative, and labor issues. Each group's decisions were input into the computer, and the impact on product and corporate performance was reported to the group, which then reviewed and revised management decisions accordingly. In this manner, during the course of the semester, each group managed their pharmaceutical firm for several years. The decisions made by each group and the performance of each corporation were independent of the other groups and their respective corporations. The performance appraisals and utility judgments in the present study were collected within the context of this simulation.

Table 1
Pharmaceutical Detailers: 1984 Quarterly Summary (Quarter 1)

Name	Repeat sales (in dollars)	New product sales activity		Expenses (in dollars)
		Contacts	% successful	
Ford	324,754	321	22	4,156
Lawrence	299,082	326	21	5,315
Walker	259,620	315	19	6,662
Clark	204,951	260	18	8,744
Robinson	341,869	372	21	4,128
Smith	247,738	313	18	7,121
Miller	222,065	335	17	8,204
Williams	284,706	307	21	5,838
Edwards	278,888	318	20	6,050
Mitchell	267,933	322	19	6,445

Procedure

Subjects were provided with the following information:

As managers, you are responsible for evaluating the performance of your employees. Attached you will find information concerning four important measures of job performance by 10 of your pharmaceutical detailers (i.e., sales representatives). The performance of these individuals is representative of the performance of your firm's entire sales force.

At the end of the fiscal year you will be asked to complete a performance appraisal concerning each of these pharmaceutical detailers. You should use the attached information in helping you to appraise the performance of these employees. Fair and accurate appraisals can be useful in several ways. These include providing feedback to the employees concerning individual performance, writing employee development plans, allocating merit increases, and determining the overall (net) value of the employee to your corporation.

The following information may be helpful to you in evaluating the contribution of the pharmaceutical detailers to your firm.

Each individual receives a base salary of $40,000. In addition, detailers receive a 6% commission on their new product sales volume. No commission is paid on repeat sales. The remaining sales (in dollars) do not represent net profit to the company. Specifically, 60% of each employee's repeat and new product sales is absorbed by production expenses, and another 20% is absorbed by overhead burden. In addition to sales of established products (i.e., repeat sales), detailers are responsible for promoting and selling the firm's new products (see the attached job description). The level of each detailer's new product sales volume can be estimated from knowledge of the number of sales contacts made by each detailer and the percentage of contacts that resulted in a successful sale of new products. It has been determined that the average amount of a new product sales contract is $625, and this dollar amount does not fluctuate significantly among detailers. Therefore, if a detailer makes 250 contacts with a success rate of 20%, the resulting sales volume can be estimated to be $31,250 (i.e., 250 × .20 × $625). Expenses generated by a sales manager (e.g., travel, meals) directly reduce the overall value of his or her profitability to the company. You should consider all of these elements in appraising each employee's performance.

Subjects were provided with information concerning repeat sales, new product sales activity, and expenses for four successive weeks (with each week providing performance information for one fiscal quarter). As an example, Table 1 contains the performance information provided to subjects on the first week (i.e., for the first fiscal quarter). In addition, each week subjects were provided with the above instructions and job description (see Table 2).

At the end of the fourth week, subjects completed the dependent measures. It should be noted that, although individual roles within the groups sometimes varied (e.g., some individuals focused on financial issues, whereas others addressed product development or marketing needs), all group members functioned at the same organizational "level" within the simulation. Thus, no effect of organizational or supervisory level could influence the subjects' SDy estimates. In addition, all ratings and estimates were completed individually and without interaction among the subjects.

Summary of Performance Information

Table 3 provides a summary of each employee's actual performance on the three job dimensions over the entire fiscal year. New product sales activity has been converted to dollar volume of sales using the formula provided to subjects in the instructions. The standard deviation of yearly repeat sales was $170,119. The standard deviation of yearly new product sales was $24,302. The standard deviation of net revenue after expenses was $43,639. (Net revenue after expenses was computed as yearly total repeat and new product sales of the employee minus production expenses, overhead burden, new product sales commission, annual salary, and travel expenses. It should be noted that all of the information required to determine net revenue after expenses was available to subjects at the time they completed the dependent measures).

Although subjects were not provided with the summary information in Table 3, they were provided with an optimal situation for providing performance appraisals

Table 2
Job Description: Pharmaceutical Detailer

Promotes use of and sells ethical drugs and other pharmaceutical products to physicians, dentists, hospitals, and retail and wholesale drug establishments, utilizing knowledge of medical practices, drugs, and medicines.

Responsibilities and Principal Activities

1. Sells established product line:
   Travels throughout assigned territory to call on regular and prospective customers to promote products and solicit orders for drugs and medicines manufactured by the company. Quotes prices and credit terms and prepares sales contracts for orders obtained. Estimates date of delivery to customer, based on knowledge of the firm's production and delivery schedules.
2. Sells new product line:
   Calls on customers, informs customer of new drugs, and explains characteristics and clinical studies conducted with drug. Discusses dosage, use, and effect of new drugs and medicinal preparations. Gives samples of new drugs to customer. Prepares sales contracts for orders obtained.
3. Controls expenses:
   Controls and minimizes expenses relative to anticipated sales volume. Keeps expense accounts for travel (automobile, airplane, etc.) and accommodations (room, meals).

and utility estimates. Specifically, each subject had performance information for each of the four quarters available at the time he or she completed the dependent measures. Second, the average interquarter correlations of employee repeat sales volume, new product sales volume, and expenses were .82, .81, and .97 respectively. Thus the performance of the employees being evaluated was relatively stable throughout the year. Third, the correlation between the employees' yearly repeat sales volume and yearly new product sales volume was 1.00. The correlation between employees' yearly repeat sales volume and yearly travel expenses was −1.00 (i.e., higher repeat sales volume was associated with better control of related expenses). The correlation between employees' yearly new product sales volume and travel expenses was also −1.00. Net revenue after expenses was also perfectly correlated with performance on each of the three job dimensions. Thus if an employee's performance on one job component was excellent (or poor), performance on the other two components was also excellent (or poor). Finally, the distributions of repeat sales, new product sales, and expenses were approximately normal (with skewness = 0.00 and kurtosis = −.42). In summary, performance appraisals and utility estimates were collected from subjects who had precise performance information available to them at the time they completed the appraisals and made their utility estimates. Also employee performance was normally distributed, stable throughout the course of the year, and perfectly correlated among the three job performance components. This perfect correlation among the job components was created to help make the performance rating task relatively easier for the subjects. Thus the ability of subjects to make accurate appraisals and utility estimates under optimal circumstances seldom encountered by managers in real organizations could be determined.

Estimates of SDy

Subjects were provided with two packages containing the dependent measures for the CREPID procedure and the Schmidt et al. (1979) procedure. The order in which subjects completed the procedures was counterbalanced.

Schmidt et al. procedure. For the Schmidt et al.

Table 3
Pharmaceutical Detailers: 1984 Annual Summary (Quarter 1 through Quarter 4)

Name	Repeat sales (in dollars)	New product sales (in dollars)	Expenses	Net revenue after expenses (in dollars)
Ford	1,299,018	185,500	17,500	228,274
Lawrence	1,196,329	170,925	21,250	201,945
Walker	1,048,458	149,756	26,650	164,007
Clark	819,804	117,156	35,000	105,363
Robinson	1,367,477	195,406	15,000	245,852
Smith	990,952	141,493	28,750	149,249
Miller	888,263	126,887	32,500	122,917
Williams	1,138,824	162,637	23,350	187,184
Edwards	1,115,548	159,387	24,200	181,224
Mitchell	1,071,733	153,100	25,800	169,981
M	1,093,641	156,225	25,000	175,600
SD	170,119	24,302	6,212	43,639

procedure, subjects were asked to estimate the total yearly repeat sales, new product sales, net sales revenue after expenses, and overall worth of an average pharmaceutical detailer. The same estimates were also obtained for employees performing at the 85th and 15th percentiles. The introduction and instructions for making these estimates were essentially identical to those used by Schmidt et al. (1979).

The CREPID procedure. The procedures described by Cascio (1982) were used to obtain the CREPID estimate of SDy. Subjects were asked to rate each of the three job components in terms of time/frequency, importance, consequence of error, and level of difficulty. Subjects had several sources of information available to assist them in making these judgments. For example, they could base their ratings on the job description and/or performance information provided to them as well as on any familiarity gained through the management simulation with the nature of such jobs in a pharmaceutical corporation. Subjects also rated each of the 10 employees on each of the three job components using the 0–200 point scale described by Cascio (1982). The remaining six steps of CREPID were completed by the authors to arrive at an estimate of SDy.

Results

The subjects appeared to take the rating and estimation tasks seriously, as these were presented in the context of the management simulation, and thus they could be perceived as potentially contributing to an individual's evaluation in the class. Although subjects had access to calculators, none were used to directly calculate the estimates. Subjects did have access to paper and pencil, but these aids were used only to complete rough calculations.

One subject's data were not included in any of the analyses because of the failure to follow directions in the CREPID procedure and inconsistent estimates (e.g., the 85th percentile was less than the average) using the Schmidt et al. (1979) procedure. Another two subjects were not included in any of the CREPID analyses because of the failure to follow directions when completing the job analysis or performance appraisal forms. Therefore, the CREPID analyses were completed using the data from 16 subjects. In addition to the one subject noted above whose data was not included in any analyses, an additional two subjects were not included in analyses of the Schmidt et al. (1979) estimates. One subject was not included because data was missing, whereas another was not included because his percentile estimates were outliers (equal to approximately 1,000% of the estimates provided by other subjects). In addition, four subjects' percentile estimates appeared to be based on quarterly rather than annual performance data. Rather than eliminate these subjects' data from subsequent analyses, it was decided to multiply the quarterly estimates times four and convert them to annual estimates of performance. Thus the Schmidt et al. (1979) estimates were also computed based on data from 16 subjects.

A Mann–Whitney test[1] was conducted to determine if any SDy estimates resulting from the CREPID or Schmidt et al. (1979) procedures differed as a result of counterbalancing the order in which subjects completed the two dependent measure packages. Results indicate that none of the SDy estimates was significantly different as a result of counterbalancing order (minimum $p > .19$). As noted above, subjects were students in a management class that had been previously separated into four groups, each of which independently managed a pharmaceutical company in a computer simulation. A Kruskal–Wallis one-way analysis of variance (ANOVA) indicated that the SDy estimates were not significantly different as a function of group membership (minimum $p > .09$).

Schmidt et al. (1979) Estimates

The difference between the 85th and average percentile estimates and the difference between the average and 15th percentile estimates provided two alternative estimates of the SDy of repeat expenses, new product sales, net revenue after sales, and overall worth. Wilcoxon matched-pairs ranked-signs tests indicated that these two alternative estimates of SDy were not significantly different (minimum $p > .17$) for any of the four parameters being estimated. The two alternative estimates of SDy were therefore averaged to obtain the final estimate of SDy for each of the four parameters. Consistent with Schmidt et al. (1979), the standard error of the mean of these estimates was also computed.

[1] Because of the potential effects of outliers and nonnormality, nonparametric tests were conducted to examine mean differences in this study.

The average estimated SDy of yearly repeat sales was $178,725. The standard error of these SDy estimates was $13,651. The actual SDy of yearly repeat sales was $170,119. This estimated SDy value was only 5% greater than the actual SDy value of repeat sales.

The average estimated SDy of yearly new product sales was $55,760. The standard error of these estimates was $27,327. The actual SDy of yearly new product sales was $24,302. Although the estimated SDy appears to be different from the actual SDy, this is the result of one of the 16 subjects whose SDy estimate of this parameter was equal to over 1,000% of the mean estimate of the remaining 15 subjects. When this single outlier estimate was removed from the computations, the resulting estimated SDy was $29,477 (with $n = 15$; the standard error of these estimates was $3,374). This adjusted estimate is still 21% greater than the actual value of SDy for new product sales.

The average estimated SDy of yearly net revenue after expenses was $119,605. The standard error of these estimates was $57,773. The actual SDy of yearly net revenue after expenses was $43,639. Again, this result was substantially affected by one outlier in the data. When this single outlier estimate was removed from the computations, the resulting estimated SDy was $65,912 (the standard error of these estimates was $17,618). This adjusted estimate is still 51% greater than the actual value of SDy for net sales revenue after expenses.

The average estimated SDy of overall worth was $83,994. The standard error of these estimates was $25,247. The actual overall worth of an employee can be a matter of definition (e.g., Cascio & Ramos, 1984; Boudreau, 1983); therefore, no actual value of the SDy of overall employee worth is offered as a contrast to the estimated SDy in this study. However, it should be noted that four of the sixteen subjects estimated the SDy of overall worth to be in the range of $200,000 to $300,000. The remaining estimates ranged from $5,000 to $75,000. The dramatic differences in the ranges of these estimates suggest that subjects may have been estimating using different interpretations of an employee's overall worth.

This point is underscored by the correlations between subjects' SDy estimates for repeat sales, new product sales, net revenue, and overall worth (see Table 4). These correlations suggest that subjects tended to consistently over- or underestimate SDy for the three components, but these individual component estimates had little relation with their overall estimates of SDy.

Table 4
Intercorrelations for Schmidt et al.—Estimates of SDy

Variable	NP	Net	Overall
Repeat sales	.69*	.69*	.26
New product sales (NP)		.97*	.00
Net			.17

Note. $N = 16$.
* $p < .01$.

CREPID Estimates

Two of the key assumptions to the CREPID technique are as follows: (a) Subjects can reliably discriminate among the elements or components of a job in terms of each component's contribution to overall worth, and (b) subjects can accurately appraise performance on each component. The results of a subjects-by-components ANOVA of the importances was used to estimate interrater reliability. The component factor was significant ($p < .01$), with interrater reliability estimated as .87.

The accuracy of the performance appraisal ratings was examined by calculating Pearson product–moment correlations between the average rating for each employee across subjects and the actual dollar value performance for employees on each component. All coefficients were |.99|, indicating that subjects could rate performance with a high degree of relative accuracy. The intercorrelations among the three sets of component ratings were also |.99|, again accurately reflecting the actual intercomponent correlations.

After applying all of the steps recommended by CREPID the estimated SDy of dollar-valued job performance was $26,485. (The standard error of these estimates was $1,381). This figure is substantially less than the SDy estimates of overall worth or net sales revenue after expenses derived from the Schmidt

et al. (1979) procedure. The mean of dollar-valued job performance was $54,313 as contrasted with the actual mean salary of $49,373 (i.e., base salary plus commission).

Discussion

The accurate estimate of SDy for repeat sales using the Schmidt et al. procedure is consistent with the Bobko et al. (1983) study in which the Schmidt et al. (1979) procedure produced estimates of SDy consistent with objective sales data. The SDy estimate of yearly new product sales was less accurate than the estimate of SDy for repeat sales but still reasonably close to the actual value. Because product sales performance information was provided in terms of sales activity rather than dollar volume of sales, subjects not only had to estimate the distribution of sales activity but also had to convert sales activity into dollars. This additional step produced a less accurate, upwardly biased estimate.

The SDy estimate of yearly net revenue after expenses was less accurate than either the repeat sales or new product sales SDy estimates. To determine point estimates of net sales revenue after expenses, subjects had to simultaneously consider seven variables (i.e., repeat sales, new product sales, travel expenses, commission, salary, production expenses, and overhead burden expenses). Boudreau (1983) noted the importance of accounting for variable costs in utility estimates. The inaccuracy and high variance of subjects' SDy estimates of net sales revenue after expenses suggests that subjects apparently have difficulty in directly estimating the effects of these variable costs.

The upward bias in SDy estimates in the present study was associated with upwardly biased point estimates of the mean of both new product sales and net revenue after expenses. As noted above, the adjusted (n = 15) estimates of SDy of new product sales and net revenue after expenses were 21% and 51% greater, respectively, than the actual SDy values of these distributions. Correspondingly, the adjusted (n = 15) estimates of the means of these two distributions were 13% and 75% greater, respectively, than the actual means of these distributions. Thus it appears that the *upward* bias in SDy estimates may have resulted from a magnitude of scale problem in which substantially overestimating the mean of a distribution resulted in a corresponding inaccuracy in the magnitude of SDy estimates. Additional research is required to determine whether *upwardly* biased SDy estimates would be found in other settings in which rater, ratee, and job characteristics are different.

The Schmidt et al. (1979) procedure was developed for use in situations in which the objective performance data needed to directly estimate SDy are not available. However, these results indicate that, as performance information becomes more difficult to convert to dollar terms, SDy estimates become less accurate and more variable. Of course, with many jobs (e.g., managers) performance information is not easily or directly translated into dollar value terms. These results suggest that estimates of SDy obtained with the Schmidt et al. (1979) procedure may be relatively inaccurate for such jobs. The Schmidt et al. (1979) estimates in the present study were most accurate where objective data could be easily converted to SDy estimates. In situations in which such objective data are available, SDy can be directly calculated, and the Schmidt et al. (1979) estimation procedure may not be necessary.

As noted above, SDy estimates seemed to suggest that subjects may have used at least two different interpretations of overall worth. The Schmidt et al. (1979) procedure asks subjects to estimate the value of "overall products and services" and to consider the cost of "having an outside firm provide these products and services." Boudreau (1983) noted that the value of products and services is not equal to the cost of obtaining them from an outside firm. Boudreau (1983) suggested that by using both of these standards in the same measure, the Schmidt et al. (1979) procedure may introduce confusion and unreliability into SDy estimates. This is further supported by the low correlations between the individual estimates and the overall estimates in Table 4. One implication of these findings is that revised instructions

that include only one of these standards (e.g., the value of overall products and services) might yield more consistent estimates of SDy.

The Schmidt et al. SDy estimate of overall worth in this study was equal to 133% of annual salary. This result is inconsistent with the Hunter and Schmidt (1983) argument that SDy can be estimated as 40% of wages. This finding may have resulted from the potential confusion in instructions noted by Boudreau (1983), from the artificial management simulation environment in which the study was conducted, or from the fact that subjects had made estimates of "net sales revenue after expenses" prior to estimating overall worth. That is, "net sales revenue after expenses" estimates may have upwardly biased estimates of overall worth.

These results are also consistent with earlier research (Bobko et al., 1983) in that the standard deviation of SDy estimates was often equal to or greater than the mean estimate of SDy. One implication of this finding is that large samples may be needed to obtain stable SDy estimates with the Schmidt et al. (1979) procedure. Estimates obtained from small samples of judges, coupled with the substantial effects of outliers noted above, could introduce considerable error into the estimates.

As noted, the SDy estimate obtained with the CREPID procedure was more conservative than the estimate of overall worth obtained with the Schmidt et al. (1979) procedure but larger than the estimate produced by the 40% rule. In addition, the standard error of the CREPID SDy estimates was much smaller than the standard error of any of the SDy estimates obtained with the Schmidt et al. (1979) procedure. Of course, these findings should not be surprising in view of the fact that, unlike the Schmidt et al. (1979) procedure, CREPID is directly tied to salary structures. In fact, it might be argued that the CREPID procedure will be dependent *only* on the job component intercorrelations and average salary.

For any given rater, SDy will be a function of the dollar value D_i assigned to a component i, the standard deviations, S_i, of performance on the 0–2.0 scale and the intercorrelations among components.

$$SDy = \left(\sum_{i=1}^{K} D_i^2 S_i^2 + \sum_{j=1}^{K} \sum_{i=1}^{K} D_i D_j r_{ij} S_i S_j\right)^{1/2} \quad (1)$$

where $i \neq j$.

Because of the scale used in CREPID all component S_is should be approximately equal.

Although Cascio states that "CREPID used magnitude estimation . . . to obtain information on performance" (Cascio, 1982, p. 170), in practice, the method calls for evaluating employees on a scale that has a linear relationship with percentiles (the scale anchors are 50 = *better than 25%*, 100 = *better than 50%*, 150 = *better than 75%*, and 200 = *better than 99%*). CREPID divides the scale values by 100 to obtain a scale with a range of 0 to 2.0. Though the resultant scale has the appearance of a magnitude estimation scale, it is not a magnitude estimation scale in terms of the judgments made by raters. As a result, all distributions should have a mean of 1.0 and a standard deviation corresponding to a rectangular distribution across the 2-point range, or S_i = .58, regardless of the actual distribution of performance. The present data support the assumption that raters use the CREPID scale in terms of percentiles. The three average S_i values for repeat sales, new product sales, and expenses were .58, .54, and .60, all close to the value expected for a rectangular distribution. Because all S_is should be approximately equal, the SDy values produced by CREPID will be a function of the D_i and r_{ij} in Equation 1.

This scaling problem would appear to represent a serious deficiency in previous SDy estimates using CREPID. However, it is a problem that can easily be solved by changing the scale and rater instructions to conform to magnitude estimation principles. One other aspect of CREPID that might be unnecessary is the decomposition of job performance into components. This step necessitates fallible judgments of component importance and employee performance on each component. Because the objective of CREPID is to arrive at a magnitude estimation scaling of employee performance, this can be done more efficiently by using a single overall magnitude estimation scaling of employees on productivity. This method would ask supervisors to assign nu-

merical values to employees based on their proportional worth in terms of productivity. Stevens (1971), noting that the numerical scale used by an individual rater is unimportant, provides a method for adjusting different subjects' judgments to a common scale. The resulting standard deviation could then be expressed as a proportion of the mean. This value could then be multiplied by average salary to obtain SDy. Such a procedure would have some similarity to the approach of Hunter and Schmidt (1983), who used empirical standard deviations of productivity measured on a ratio scale. If the assumption is made that a magnitude estimation scaling of employees on overall productivity represents a similar ratio scaling, then the relationship between a revised CREPID procedure and Hunter and Schmidt is SDy(CREPID) = ½ SDy(Hunter–Schmidt), because Hunter and Schmidt assume that salary represents one half of the cost of production. It should be noted that Boudreau (1983) argued that the appropriate economic value for analyzing utility is not wages but the value of what is produced, adjusted for variable costs, taxes, and discounting.

Conclusion

The present study examined the methods and assumptions of two procedures for estimating dollar standard deviations of performance. Using the Schmidt et al. procedure, judges in a simulated environment accurately estimated SDy when dollar values were direct and explicit, consistent with earlier research. However, the same judges were unable to provide accurate estimates of SDy as the translation of performance to dollars became more complex. Although the simulated setting provided relatively optimal conditions in terms of the amount and type of information available, the results are necessarily limited in generalizing to actual job settings. It may be that experienced supervisors can more accurately translate the performance of their employees into dollar values even under complex conditions. The present results suggest that different judges may have interpreted the Schmidt et al. instructions differently. More data from field settings are needed to determine whether training, changes in directions for judges, or both will facilitate more accurate SDy estimates.

The CREPID approach was shown to provide reliable assignments of job component importance and accurate estimates of relative employee performance. Although this consistency between subjective ratings and objective data is encouraging, it should be remembered that subjects were, in a sense, simply reporting back data provided to them under optimal conditions. The ability of subjects to make accurate CREPID appraisal ratings under other conditions (e.g., where component intercorrelation is low or where subordinates and supervisors have ongoing interaction) needs to be investigated.

It was demonstrated that the scale used in CREPID should yield a rectangular distribution with a fixed standard deviation. As an alternative, direct magnitude estimation of employee performance could be used to estimate SDy for complex jobs, such as that performed by managers. If the economic assumptions of Hunter and Schmidt (1983) hold for such jobs, then SDy should be estimated by multiplying the magnitude estimation SD by twice the average annual salary.

The importance of accurate utility estimates for industrial/organizational psychology is obvious. Further effort is needed to examine the ability of judges to estimate SDy when actual values are known to investigators, and the translation into dollars is complex. Such research is the best way to provide evidence that utility estimation procedures are useful and accurate for situations in which objective data are unavailable. Attempts should also be made to take into account the issues raised by Boudreau (1983) as they affect the net dollar worth of employee intervention programs and the suggestions of Bobko et al. (1983) regarding the use of sequential judgment and feedback for improving judge's estimates (see, e.g., Burke & Frederick, 1984).

References

Bobko, P., Karren, R., & Parkington, J. J. (1983). Estimation of standard deviations in utility analyses: An empirical test. *Journal of Applied Psychology, 68,* 170–176.

Boudreau, J. W. (1983). Economic considerations in estimating the utility of human resource productivity improvement programs. *Personnel Psychology, 36,* 551–576.

Brogden, H. E. (1949). When testing pays off. *Personnel Psychology, 2,* 171–185.

Burke, M. J., & Frederick, J. T. (1984). Two modified procedures for estimating standard deviations in utility analysis. *Journal of Applied Psychology, 69,* 482–489.

Cascio, W. F. (1982). *Costing human resources: The financial impact of behavior in organizations.* Boston: Kent.

Cascio, W. F., & Ramos, R. A. (1984). *The economic impact of assessment center technology on management performance.* Manuscript submitted for publication.

Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions* (2nd ed.). Chicago: University of Illinois Press.

Hunter, J. E., & Schmidt, F. L. (1983). Quantifying the effects of psychological interventions on employee job performance and work-force productivity. *American Psychologist, 38,* 473–478.

Schmidt, F. L., Hunter, J. E., McKenzie, R. C., & Muldrow, T. W. (1979). Impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology, 64,* 609–626.

Stevens, S. S. (1971). Issues in psychophysical measurement. *Psychological Review, 78,* 426–450.

Weekley, J. S., Frank, B., O'Connor, E. J., & Peters, L. H. (1983). *A comparison of three methods of estimating the standard deviation of performance in dollars.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles, CA.

Received November 8, 1984
Revision received January 25, 1985 ■