

RELATIONSHIP BETWEEN SEX, RACE, AND JOB PERFORMANCE LEVEL AND THE GENERATION OF CRITICAL INCIDENTS¹

Michael G. Aamodt
Wilson W. Kimbrough
Robert J. Keller
Krista J. Crawford
University of Arkansas

Ninety-six subjects participated in a study in which the relationships between sex, race, and job performance level and the generation of critical incidents were examined. The results indicated that race was the only variable that affected the critical incidents. These results were discussed in terms of future applications.

Job analysis is the cornerstone of personnel selection (Cascio, 1980) and the Critical Incident Technique is one of the major job analysis methods (Levine, Bennett, & Ash, 1979). Even though the importance of job analysis is unquestioned, little research has been conducted on many of the variables affecting the outcomes of various job analysis methods (Levine, Ash, & Bennett, 1980). Thus, it is the purpose of this study to describe previous research concerning variables affecting the outcomes of the critical incident method of job analysis and furthermore, to investigate the relationship of sex, race and job performance level on the generation of critical incidents.

Even though similar methods were used prior to the 1950's, the Critical Incident Technique was formally developed and reported by Flanagan (1954). Basically, this technique involves a

¹Requests for reprints should be sent to Michael G. Aamodt or Wilson W. Kimbrough, Department of Psychology, University of Arkansas, Fayetteville, AR 72701.

systematic set of procedures for collecting direct observations of human behavior. These observations, called critical incidents, involve behavior which has made the difference between success and failure in the carrying out of a job (Flanagan, 1954). These incidents are usually obtained by asking a job incumbent to record incidents of job behavior that appear to be critical to job performance.

One of the first investigations of the technique itself was conducted by Flanagan, Miller, Burns, Hendrix, Stewart, Preston, and West (1953). In this study, Flanagan et al. (1953) discovered that better critical incident recording results can be obtained when incidents are recorded daily rather than weekly. Flanagan et al. (1953) also found that the number of incidents was similar regardless of whether the incidents were obtained through interview or through daily performance records, and Wagner (1948) found that incidents obtained from group interviews are of the same quality as those obtained through individual interviews.

Along these lines, Wagner (1950) found that the source of the incidents (patients, dentists, or dental instructors) had an effect upon the incident categories. In this study, patients reported more relationship-related incidents and fewer technical-proficiency incidents than did the dentists or dental instructors. Similar results were found by Smit (1952) when she investigated psychology instructors.

Finkle (1950) found that the types of incidents obtained do not greatly change as a result of the wording of the instructions and that the order in which the incidents are asked for (effective behavior vs. ineffective behavior) also does not make a significant difference. These findings were supported by Andersson and Nilsson (1964).

More recently, Levine, Ash, and Bennett (1980) compared the Critical Incident Technique with job elements, the Position Analysis Questionnaire, and task analysis. Levine et al. (1980) found that personnel selection plans that used the Critical Incident Technique were rated by users to be of a higher quality than the plans resulting from the other methods. However, the critical incident technique was also more expensive to use than were the job elements approach or the Position Analysis Questionnaire.

The results of Levine, Ash, and Bennett (1980) are interesting when compared to the results of Levine, Bennett, and Ash (1979) who found that even though the Critical Incident Technique was rated as producing better results, it was used significantly less than either job elements or task analysis. Moreover, fewer personnel

selection practitioners had received formal training in the critical incident technique than had received formal training in job elements or task analysis. However, with the increased interest in Behaviorally Anchored Rating Scales, the critical incident technique remains an important tool in a personnel researchers armamentarium.

It is the purpose of this present study to investigate one of the neglected components of the critical incident process, the individuals supplying the incidents. It is common practice to have either all employees or randomly selected employees supply the incidents. While this is not necessarily a bad practice, it is possible that certain variables may be operative which might influence the types of incidents reported. The variables selected for investigation in this study were sex, race, and the job performance of the individual. These particular variables were chosen because not only have they been found to be important in previous personnel research (Arvey, Passino, Lounsbury, 1977; Bigoness, 1976), but also because it is conceivable that people who differ in job characteristics may perceive their job in different ways.

Method

Subjects. The subjects in the study were 96 Resident Assistants, Head Residents, and Assistant Head Residents who had either full or partial responsibility for one of ten residence halls at the University of Arkansas. Information on the sex, race, and job performance level of each Resident Assistant and Assistant Head Resident was obtained from the Head Residents. The job performance level of each Resident Assistant was assessed through a nine-point graphic rating scale and was made by the respective Head Resident. Resident Assistants receiving ratings of less than seven were designated as being lower in performance while Resident Assistants receiving ratings of seven or more were designated as being higher in performance.

Procedure. The following procedure was used in the current study:

1. Obtaining Critical Incidents—Each subject in the study was presented with a form in which they were given the following instructions: "Think of the best Resident Assistant that you have ever known. Now describe in detail one incident that reflects why this person was the best." A similar instructional set was utilized for the incident which typified the worst Resident Assistant.

2. Incident Segmentation—Three of the authors cooperatively

broke down each of the participants incident reports into one or more specific behavioral incidents. This procedure resulted in 312 separate examples of effective or ineffective Resident Assistant behavior.

3. Category Formation—The 312 incidents were independently sorted by three of the authors into an unspecified number of categories. Discrepancies regarding the number and definitions of the categories were resolved in a group deliberation among the judges. The outcome of the deliberations resulted in the following mutually agreed upon categories: availability, responsibility, fairness, discipline, self-confidence, interest in residents, authoritarianism, social skills, programming, self-control, confidentiality, self-confidence, and self-adherence to the rules. The categories and definitions were then submitted to the eight Head Residents for confirmation. This procedure resulted in slight changes in the category definitions but no additions or deletions in the number of categories.

TABLE 1
Distribution of Critical Incidents by Sex

Category	Male (<i>n</i> =50)	Female (<i>n</i> =46)
Interest in residents	20.30	14.11
Availability	12.03	15.34
Responsibility	10.53	11.04
Fairness	9.02	9.82
Self-adherence to rules	7.52	11.04
Social Skills	9.77	7.98
Programming	6.77	6.75
Self-confidence	5.26	7.98
Rule enforcement	6.77	6.75
Total number of incidents	133	163

4. Incident Sorting—The 312 incidents were again independently sorted into the 13 categories by three Resident Assistants. If two or more (67%) agreed that a particular incident belonged in a category, the incident was retained. The percentage of agreement for at least two judges across the 13 categories was 97.8, and the procedure resulted in the discarding of six incidents. This level of reliability was considered high enough to have confidence in the judgments of the raters and in the selection of categories.

Results

Chi-square tests were used to determine if the frequency of incidents per category differed between any of the groups. Only categories with expected values of five or greater were retained (Hays, 1981). The categories of self-control and confidentiality were dropped for all comparisons because they did not meet this criteria. Thus, the number of categories will differ across the 3 comparisons. As indicated in Table 1, the results show that the

TABLE 2
Distribution of Critical Incidents by Race

Category	White % (<i>n</i> = 71)	Black % (<i>n</i> = 25)
Interest in residents	16.23 (25) ^a	5.88 (3)
Availability	16.23 (25)	13.73 (7)
Responsibility	12.33 (19)	13.73 (7)
Fairness	16.23 (25)	5.88 (3)
Self-adherence to rules	21.43 (33)	33.33 (17)
Social Skills	17.53 (27)	27.45 (14)
Total number of incidents	154	51

^aRepresents number of incidents.

frequency of incidents in each category was not moderated by the sex of the person generating the incidents, $\bar{\chi}^2(8)=8.42$, *n.s.*, thus, males generated the same types of incidents as did females.

However, as shown in Table 2, the type of incident was influenced by the race of the individual generating the incidents. That is, whites generated different frequencies of the various categories of incidents than did blacks, $\bar{\chi}^2(5)=12.73$, $p<.05$. Only six categories were retained for this analysis as seven of the categories had expected values of less than five. It should also be noted that there was no significant difference between the job performance ratings of whites and those of blacks.

TABLE 3

Distribution of Critical Incidents by Job Performance Level^a

Category	High Level %	Low Level %
Interest in residents	13.97 (19) ^b	16.80 (21)
Availability	16.18 (22)	10.40 (13)
Responsibility	10.29 (14)	12.00 (15)
Fairness	6.62 (9)	11.20 (14)
Self-adherence to rules	11.76 (16)	9.60 (12)
Social Skills	9.56 (13)	8.00 (10)
Programming	7.35 (10)	5.60 (7)
Self-confidence	6.62 (9)	6.40 (8)
Rule enforcement	8.09 (11)	5.60 (7)
Authoritarianism	5.15 (7)	7.20 (9)
Counseling skills	4.41 (6)	7.20 (9)
Total number of incidents	136	125

^a The number of incidents for this comparison will not equal the number for the sex comparison due to some subjects not being rated on their performance.

^b Represents number of incidents

Overall, as shown in Table 3, the job performance level of the individual did not moderate the type of incident categories generated by the subjects, $\chi^2(10)=6.37$, *n.s.*.

Discussion

As indicated in the results, the types of incidents generated by males were similar to those generated by females. Thus, males and females seem to observe the same types of incidents and consider similar behaviors to be critical to job performance.

At first glance, it should appear that job performance also does not influence the type of critical incidents that are generated by employees. In the present data, the overall difference in the incident distributions was not significant; and only one of the eleven individual categories showed a significant difference in the frequency of incidents in the category. However, in the current study, high vs. low performing individuals were determined by a median split. Because the mean rating was near seven, the low performing individuals may have only been relatively lower performing individuals and not unsatisfactory individuals. If this study was replicated in a setting with a greater percentage of low performing individuals, then job performance may show an effect on the type of incidents that are generated. Until then, the influence of job performance on critical incidents is still unclear.

The major finding of this study is that the frequency of incidents in each category that were generated by Whites were not the same as those that were generated by Blacks. The results of the current study also suggest that the practice of having employees create incidents for specified dimensions (e.g. Bernardin, 1979) rather than deriving the categories from the incidents, serves not only to obtain an adequate number of incidents per category, but also, may reduce any influence that race might play in the job analysis. Furthermore, the actual effect on a job analysis of differences in incident frequencies by various racial subgroups is a function of how the critical incidents are used. If the frequency of incidents in a particular category is used to weight the importance of the category to overall job performance, then the results of the job analysis may be affected by the race of the individuals generating the incidents. However, if the frequencies are ignored, the actual category types will probably not differ.

Overall, the results of this study seem to suggest that critical incidents are only slightly influenced by the characteristics of the

subgroups which generate them. The main moderator in incident generation seems to be race, but more investigation is necessary on the other two variables studied in this investigation and other demographic and personality variables which may affect the generation of critical incidents groups. Furthermore, because of the moderate number of Blacks in the current study and because of the lack of a theoretical explanation for our findings, further research on groups with larger numbers of Blacks is necessary.

REFERENCES

- ANDERSSON, B. E., & NILSSON, S. G. Studies in the reliability and validity of the critical incident technique. *Journal of Applied Psychology*, 1964, 48, 398-403.
- ARVEY, R. D., PASSINO, E. M., & LOUNSBURY, J. W. Job analysis results as affected by sex of incumbent and sex of analyst. *Journal of Applied Psychology*, 1977, 62, 411-416.
- BERNARDIN, H. J. The predictability of discrepancy measures of role constructs. *Personnel Psychology*, 1979, 32, 139-153.
- BIGONESS, W. J. Effect of applicant's sex, race, and performance on employers' performance ratings: Some additional findings. *Journal of Applied Psychology*, 1976, 61, 80-84.
- CASCIO, W. F. *Applied psychology in personnel management*. Reston, VA: Reston Publishing Company, 1978.
- FINKLE, R. B. A study of the critical requirements of foremanship. *University of Pittsburgh Bulletin*, 1950, 46, 291-297.
- FLANAGAN, J. C. The critical incident technique. *Psychological Bulletin*, 1954, 51, 327-358.
- FLANAGAN, J. C., MILLER, R. B., BURNS, R. K., HENDRIX, A. A., STEWART, B., PRESTON, H. O., & WEST, E. D. *The performance record for foreman and supervisors*. Chicago, IL: Science Research Associates, 1953.
- HAYS, W. L. *Statistics*. New York, NY: Holt, Rinehart, and Winston, 1981.
- LEVINE, E. L., BENNETT, N., & ASH, R. A. Evaluation and use of four job analysis methods for personnel selection. *Public Personnel Management*, 1979, 8, 146-151.
- LEVINE, E. L., ASH, R. A., & BENNETT, N. Exploratory comparative study of four job analysis methods. *Journal of Applied Psychology*, 1980, 64, 524-535.
- SMIT, J. A. A study of the critical requirements for instructors of general psychology courses. *University of Pittsburgh Bulletin*, 1952, 48, 279-284.
- WAGNER, R. F. A group situation compared with individual interviews for securing personnel information. *Personnel Psychology*, 1948, 1, 93-107.
- WAGNER, R. F. A study of the critical requirements for dentists. *University of Pittsburgh Bulletin*, 1950, 46, 331-339.