

Selecting for Retention:

Predicting Turnover Using Alternative Analytic Techniques

Michael D. Biderman
University of Tennessee at Chattanooga

H. Kristl Davison
University of Hartford

Kay K. Swartout
McKee Foods Corporation

Mark Newsome
McKee Foods Corporation

Authors' Note: We thank Chrisann Lee and Meg Cole for their invaluable help in data management. Correspondence regarding this article should be sent to Michael Biderman, Department of Psychology / 2803, U.T. Chattanooga, 615 McCallie Ave., Chattanooga, TN 37403. E-mail: Michael-Biderman@utc.edu

Paper presented at the 18th Annual Society for Industrial and Organizational Psychology Conference, Orlando, FL, 2003.

ABSTRACT

Most theories of turnover in organizations have concentrated on factors that can only be measured after employees have been hired, even though much research has suggested that factors available prior to employment may predict turnover. This research investigated several of those pre-employment factors. Results of three different analytic techniques suggested that applicant employment history, whether applicants had friends at the organization, and interview ratings were reliable predictors of turnover, as was a test of manual dexterity. Implications of the results for current theories of the turnover process are discussed.

Most of the current theories of turnover in organizations have emphasized attitudinal factors that become salient after employees have been hired. For example, Hom and Griffeth's (1991) model places job satisfaction as the beginning state in the process leading to turnover. However their model has little to say about the etiology of the differences in satisfaction that lead to differences in turnover. In comparison, Mitchell, Holtom, and Lee (2001) acknowledged that an accumulation of job dissatisfaction was one possible beginning of the turnover process but proposed that so-called "shocks" – unexpected events that cause employees to begin thinking about turning over – might initiate turnover processes in lieu of job satisfaction. However, in their model, as in Hom and Griffeth's (1991) model, the factors primarily responsible for the initiation of the turnover process are factors that, like job dissatisfaction, are assumed to occur *after* employee selection.

While it may be true that job satisfaction is the proximal initiator of the processes that lead to turnover, there are both theoretical and empirical reasons for turnover researchers to investigate factors that occur prior to employment. First, there is mounting evidence linking job satisfaction to dispositions that exist prior to employment. This evidence suggests that job satisfaction may very well be the consequence of factors present at time of selection. Secondly, the predictors of turnover at the time of selection are those that selection specialists can use most effectively to manage turnover.

The focus of the research presented here is on factors measurable at time of selection. Although there have been many studies of the efficacy of such factors, recent meta-analyses have found little evidence for the usefulness of many of them (Cotton & Tuttle, 1986; Griffeth, Hom, & Gaertner, 2000; Hom & Griffeth, 1995; Maertz & Campion, 1998). This is perhaps best exemplified by the conclusion of Griffeth, et al. (2000) that "this new meta-analysis replicated the previous findings for most demographic predictors, affirming their modest predictive strength." (p. 479). Some reviews, however, have found evidence for the usefulness of predictors measurable at time of selection, including weighted application blanks (Breaugh & Dosset, 1989) and biographical data (Schmitt, Gooding, Noe, & Kirsch, 1984). Other recent research has provided evidence that personality traits (Boudreau, Boswell, Judge, & Bretz, 2001; Parks & Waldo, 1999), cognitive ability (Boudreau et al., 2001), and interview performance (McDaniel, Whetzel, Schmidt, & Maurer, 1994; Schmidt & Rader, 1999) are predictive of turnover. Thus,

although factors available at selection play small roles in current models of turnover, there is evidence that such factors are worthy of further investigation.

A second purpose of the present research is to illustrate the similarities and differences between conclusions drawn using different analytic techniques. The majority of turnover research has used a dichotomous turnover variable as the dependent variable and examined its relationship with the various predictors using linear regression/correlation techniques. Some, responding to suggestions by Williams (1990) have computed the correlations adjusting the overall turnover rates so that they were equivalent to 50% turnover. Most of these studies have relied on and reported linear correlation techniques, (i.e., Pearson r). There are two separate criticisms of such an analytic technique. First, it is now well known that logistic regression techniques are more appropriate than those based on the linear model for studying relationships involving dichotomous dependent variables. For situations in which the turnover rates are not too deviant from .5, there may be little difference between conclusions arrived at via linear correlation and those obtained via logistic regression (Cox & Wermuth, 1992), but when turnover rates are very small or very large, the conclusions based on the two methods may diverge.

The second criticism applies to the treatment of the outcome as a dichotomy in both linear and logistic regression methods. Such treatment ignores differences in tenure associated with turnover. For example, Figure 1 presents two groups in which the turnover rates were comparable within a window of time, yet the distributions of employment tenure until turnover were quite different. This suggests that the turnover dichotomy alone is an imperfect indicator of the tenure/turnover construct. A solution to this problem is the use of survival analysis as the analytic technique for studying turnover. Long employed in medicine to study the relationship of length of survival to type of treatment, survival analysis was specifically designed to deal with data in which the dependent variable is a construct represented by both duration until an event and the occurrence/non occurrence of the event, and in which some study participants might never reach the event (Wright, 2000). This latter possibility, termed right censoring, is very common in turnover research, as many employees within organizations *do not* turn over during the study's timeframe. Because survival analysis takes both the duration until turnover and the event of turnover into account, it is a more appropriate technique than either linear or logistic regression analysis for such research.

The study reported here investigated the relationship of turnover to several potential predictors, all of which would typically be available at time of selection. In addition, we compared the three common methods of analysis of such data - linear regression, logistic regression, and survival analysis employing Cox regression.

Method

Participants

As part of a consulting project on turnover, the authors obtained the employment records for 890 employees hired into entry level manufacturing positions. The positions included production, shipping, and environmental services. The participants represented all employees hired at a single company facility between February 1, 1999 and January 3, 2001. Of the participants, 57% were female, 67% were White, 28% African American/Black, and 7% Other, and 42% were Married, 47% Single, and 9% Divorced. Mean age was 31.2 (SD = 9.58). At the end of the observation window, 38.7% of the employees had left the company, 32.8% voluntarily and 5.8% involuntarily.

Procedure

The variables chosen for inclusion in the study were selected from those available in the firm's HRIS database and from application blanks and interviewer sheets. The information from the application blanks and interviewer sheets were entered into Excel spreadsheets and merged with the information from the HRIS database.

Measures

Predictors. The choice of predictors was made based partly on consideration of the turnover literature and partly on what was part of the organization's practice. Traditional demographic predictors used were sex of the applicant, ethnic group (contrast coded as White or African-American/Black vs. Other, and White vs. African-American/Black), and age in years at the time of application. Application-related predictors included whether the applicant received his or her preferred shift (yes/no), whether the applicant had friends at the company (yes/no), the applicant's wage rate upon employment, and the difference in dollars between the applicant's wage upon employment and wage at the most recent prior employment. Because data were missing for a substantial number of cases for the variable concerning whether the applicant had friends at the company and for the wage differential variable, two separate variables were created to represent the "missingness" of these variables. These variables allowed us to use as much data

as possible and to minimize problems associated with what might be called “wandering sample characteristics” as variables with different patterns of missing values are analyzed. These “missingness” variables were included in and controlled for in all analyses involving their corresponding variables (see Cohen & Cohen, 1983, Ch. 7).

Information was available on number of previous employers in the past five years and on length of time at the most recent employment. Because these two variables were negatively correlated ($r = -.55$), an index combining the two was created, by computing Z-scores for the nonmissing values of the two variables. The index was then computed as the difference between the Z of number of days at the most recent employment minus the Z of the number of past employments. This index thus served as an indicator of the quality of the applicant’s employment history, with positive values representing a better history. A “missingness” variable was also created for this measure, and was controlled for in all analyses involving this index.

Two additional measures from the selection process were included. First, all applicants had been given a basic manual dexterity test involving moving pegs from one board to another, and the time taken to complete this task was recorded. Secondly, a rating of applicant appearance was available from the records of eight company interviewers. Preliminary analyses of the interviewer ratings suggested there were individual differences in the average ratings given by the eight interviewers. To account for these individual differences, the interviewer’s mean was subtracted from the raw rating to create the appearance rating measure. A missingness variable for the appearance rating measure also was created and controlled.

Turnover. Of the 344 persons who left during the study window, 52 were recorded as involuntary turnover, initiated by the company. The remaining 292 were recorded as voluntary turnovers. Only the voluntary turnovers were recorded as turnovers in the analyses that follow, and the data of the 52 persons released by the company were treated as right-censored cases in the survival analyses. (Results did not change when the 52 involuntary leavers were excluded from all analyses.) In addition, the tenure of each employee, up to the end of the timeline, was recorded

Analyses

Three parallel sets of analyses of the data were performed. Linear regression/ correlation analyses, logistic regression analyses, and survival analyses using Cox proportional hazards regression were conducted. Cox proportional hazards regression was used because it is

empirically robust, requiring only the assumption that the hazard function for all participants be a constant multiple or proportion of an overall hazard function whose form need not be specified (Morita, Lee, & Mowday, 1993).

Both single-predictor and multiple-predictor analyses involving all the predictors were performed. The single-predictor analyses were performed on the assumption that practitioners might desire information on the predictive ability of a variable in the absence of controls for other variables. Multivariate tests were performed for the theoretical value of knowing the relationship of turnover to the unique variation in a variable after having partialled out the effects of the other variables.

Because of concerns that the types of jobs represented in the sample might be unique to the particular manufacturing facility and as there were differences in physical demands of different jobs, we grouped employees into job categories representing similar tasks, created group-coding variables representing these categories, and partialled out the job categories in all the analyses reported below. Moreover, as mentioned above, for those variables with missing values, the missingness of the variable was also partialled out in the analyses, yielding a result controlling for job category and representing the relationship among those for whom the variable was not missing. Thus, the results labeled single-variable are results in which both job category and missingness have been partialled out. Those labeled multi-variable also partialled out the other predictors and their missingness variables.

Results

Table 1 presents the means, standard deviations, and correlations of the study variables. The single-predictor analyses, controlling only for job category and missingness (where appropriate) are presented in the left-hand columns of Table 2. As can be seen from inspection of the table, with one exception, the results are uniform across type of analysis. Sex and age were not predictive of turnover. Those who indicated they were White or African-American/Black had significantly higher turnover rates than those indicating otherwise, but White vs. African-American/Black did not predict turnover.

Whether an applicant received his or her preferred shift did not predict turnover, but applicants who had friends at the organization had lower turnover rates than those who indicated no friends. The wage difference variable also failed to predict turnover across all three analytic techniques and are discussed below. The applicant's score on the employment history index was

predictive of turnover, with those scoring higher (longer time at last employment, fewer employments in last five years) having lower turnover. Those who took longer to complete the dexterity test also had lower turnover rates, as did those applicants whose appearance ratings were greater than the interviewer's mean.

The single-predictor linear and logistic regression analyses gave almost identical results across all variables. With one exception, the survival analysis results also paralleled the linear and logistic analyses. The notable exception was in the analysis of wage rate. Both the linear and logistic analyses suggested that the relationship is significantly negative, with employees who receive high wages having lower turnover rates. The survival analysis results, based on both tenure and turnover, suggested that relationship to wage rate is negligible.

The multiple-predictor results are presented in the rightmost columns of Table 2. For these analyses, all variables – those representing job categories, missingness, and the variables representing the other predictors – were entered simultaneously, so each result controls for all others in the analyses. For the most part, the multi-variable results are similar to the single-variable results, although, of course, the interpretation is different. As with the single-predictor results, the multiple-predictor results suggest that after partialling out the other variables, those classified as White or African-American/ Black have lower turnover than those classified as Other, applicants with friends will have lower turnover rates than those without, and those with a good employment history will have lower turnover. The multi-variable analyses also paralleled the single-variable analyses of both time to complete the dexterity test (longer times were associated with lower turnover) and appearance rating relative to the interviewer's mean, with higher ratings associated with lower turnover. The contradictory results concerning wage rate were also obtained in the multiple-predictor analyses, with linear and logistic regression suggesting a negative relationship and survival analysis suggesting a negligible relationship.

Discussion

The current study was designed to examine several pre-employment predictors of turnover using three different statistical analytic techniques. As most prior studies of turnover have simply used Pearson correlations or linear regression, the current study shows the relationship between results of the three analytic techniques and serves to help bridge the gap between results obtained from studies employing the traditional methods (mainly linear regression/correlation) and those obtained using more appropriate methods (e.g., survival

analysis). The comparisons of the three types of analyses indicated that for these data, there were no substantial differences in results between the linear and logistic regression analyses of turnover outcomes. Since the base rate of turnover for these data was greater than .3, this finding is not unexpected since the largest differences between linear and logistic regression analyses are found when the base rate is far from .5.

The differences between the event history analysis and the linear and logistic regression analyses were substantial in the case of one variable - wage rate. Analyses of turnover only, based on either linear or logistic regression, would have led one to conclude that turnover is significantly negatively related to wage rate. The survival analysis, based on *both* tenure and turnover outcomes, gave a very different picture, suggesting that the composite criterion is not related to wage rate. Given the growing acceptance of survival analysis for such data, this result serves as a caution to those relying exclusively on either linear or logistic regression analyses. It also suggests that studies of the turnover process might profit from separate examination of tenure and turnover.

With regard to the specific findings of the analyses, our results are in keeping with recent meta-analytic results (Griffeth, et. al., 2000). For example, we found that sex of the applicant was not related to turnover even when controlling for other predictors used in the present study. We also found no differences in turnover between applications recorded as White and those recorded as African American/Black, although those recorded as Other had significantly lower turnover rates than either of the two major racial groups. The Other category for this study comprised 40 persons whose ethnic groups were identified as Asian American, American Indian, and Hispanic. Of these 24 were Hispanic. This result may reflect the fact that those in the Other category of our study had the most to lose by turning over and the greatest motivation to stay.

We found no relationship of turnover to age of the applicant, in contrast to the findings of Hom and Griffeth (1995), who concluded that “older employees with long tenure in a company quit less often than younger and short-tenure employees did” (p. 39). However, since our sample consisted of only hires within the two-year study window, tenure differences in our sample were small. This may account for the difference between our results and those reported by Hom and Griffeth.

We included a collection of predictors that might be considered to be formative indicators of job satisfaction – whether the applicant received a preferred shift, whether the applicant had

friends at the company, wage rate, and the difference between the current and past wage. Of these only one – whether the applicant had friends at the company – was related to turnover. The importance of friends has been noted by Buckingham and Coffman (1999). The nature of this relationship is what would have been expected if having friends at the company leads to greater job satisfaction. . However, in keeping with the meta-analytic results of Griffeth and Hom (2000), neither wage rate nor the difference in wage rates between the current and most previous job was a predictor of turnover. Finally, whether or not the applicant received the preferred shift did not predict turnover. Thus, three variables that would be logical candidates for formative indicators of job satisfaction failed to materialize as useful predictors of turnover.

In keeping with the findings of Judge and Watanabe (1995), the index of employment history based upon the length of time at the previous job minus the number of previous jobs was a significant predictor of turnover. This finding extends those of Judge and Watanabe to the population of employees in manufacturing and suggests that indeed, the past is prologue, even among applicants to blue-collar manufacturing jobs. This result is consistent with those found for other types of biographical information such as that obtained in weighted application blanks, scores on which have been found to be highly related to turnover (Griffeth & Hom, 2000; Hom & Griffeth, 1995).

The results of the appearance rating score from the pre-employment interview also correspond nicely with the two recent meta-analyses involving interview data (McDaniel et. al., 1994; Schmidt and Rader, 1999). Those with higher appearance rating scores had lower turnover rates. This result was consistent across all three analyses.

The most curious result is that involving the dexterity test scores. We found that persons who completed the test slowly had lower turnover rates than those who completed the test quickly. We have found no previous research relating dexterity to turnover. Given the assumption that those who completed the test quickly would be more likely to succeed at jobs for which dexterity was important, we would expect quicker applicants to have better job performance, and, based on the finding that performance is slightly inversely related to turnover (Williams & Livingstone, 1994), we expected the quickest applicants to have the lowest turnover rates. As it was the policy of the company to place applicants based on dexterity scores, it was possible that the dexterity effect might be explained by job placement, but the fact that the effect was found after controlling for type of job argues against that explanation. One final possible

explanation is that the quickest applicants were those with greater flexibility in the job market, and were more likely to engage in job search behavior than those who had high scores on the dexterity tests. Unfortunately, we have no data to test this hypothesis.

Implications and Conclusion

We have identified several variables measurable at time of selection that are related to turnover. A logical next step, given the importance of job satisfaction in current models of turnover, is to attempt to link these variables to job satisfaction to assess the extent to which it serves as a mediator of the relationships discovered here. For example, the extent to which job satisfaction mediates the relationship between turnover and whether friends are already employed at an organization is an empirical question, and the status of this variable as formative indicator of job satisfaction should be tested. It is also possible that employment history might be best viewed in a causal relationship with job satisfaction. As Judge and Watanabe (1995) suggested, dispositional factors resulting in differences in job satisfaction may play a key role here. It may be that some employees are dispositionally dissatisfied and may move from job to job in response to that dissatisfaction. Finally, the relationship, if any, between dexterity and job satisfaction should be investigated.

By examining pre-employment predictors of turnover using three different methods of statistical analysis, the current study has provided more appropriate tests of the relationship between turnover and some demographic variables currently explored in the literature, along with additional formative predictors of turnover. Although, for the most part, similar results were obtained across the three analysis techniques, researchers are encouraged to carefully conceptualize whether they are measuring tenure, turnover, or both, and select the statistical technique appropriate to that conceptualization.

References

- Boudreau, J. W., Boswell, W. R., Judge, T. A., & Bretz, R. D., Jr. (2001). Personality and cognitive ability as predictors of job search among employed managers. *Personnel Psychology, 54*, 25-50.
- Breaugh, J.A., & Dosset, D. L. (1989). Rethinking the use of personal history information: The value of theory-based biodata for predicting turnover. *Journal of Business and Psychology, 3*, 371-385.
- Buckingham, M. & Coffman, C. (1999). *First, Break all the Rules*. New York: Simon & Schuster.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cotton, J. L., & Tuttle, J. M. (1986). Employee turnover: A meta-analysis and review with implications for research. *Academy of Management Review, 11*(1), 55-70.
- Cox, D. R., & Wermuth, N. (1992). A comment on the coefficient of determination for binary responses. *American Statistician, 46*, 1-4.
- Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management, 26*(3), 463-488.
- Hom, P. W., & Griffeth, R. W. (1991). A structural equations modeling test of a turnover theory: Cross-sectional and longitudinal analysis. *Journal of Applied Psychology, 76*, 350-366.
- Hom, P. W., & Griffeth, R. W. (1995). *Employee turnover*. Cincinnati, OH: South-Western College Publishing.
- Hom, P. W., & Kinicki, A. J. (2001). Toward a greater understanding of how dissatisfaction drives employee turnover. *Academy of Management Journal, 44*, 975-987.
- Judge, T. A., & Watanabe, S. (1995). Is the past prologue?: A test of Ghiselli's hobo syndrome. *Journal of Management, 21*(2), 211-229.
- Maertz, C. P., Jr. & Campion, M. A. (1998). 25 years of voluntary turnover research: A review and critique. *International Review of Industrial and Organizational Psychology, 13*, 49-81.

- McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology, 79*(4) 599-616.
- Mitchell, T. R., Holtom, B. C., & Lee, T. W. (2001). How to keep your best employees: Developing an effective retention policy. *Academy of Management Executive, 15*(4), 96-108.
- Morita, J. G., Lee, T. W., & Mowday, R. T. (1993). The regression-analog to survival analysis: A selected application to turnover research. *Academy of Management Journal, 38*(6) 1430-1464.
- Parks, C. A., & Waldo, D. (1999). Assessing voluntary turnover likelihood using personality traits measured during pre-employment selection. *Current Research in Social Psychology, 4*, 135-145.
- Schmidt, F. L., & Rader, M. (1999). Exploring the boundary conditions for interview validity: Meta-analytic validity findings for a new interview type. *Personnel Psychology, 52*, 445-464.
- Schmitt, N., Gooding, R. Z., Noe, R. D., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology, 37*, 407-422.
- Williams, C. R. (1990). Deciding, when, how, and if to correct turnover correlations. *Journal of Applied Psychology, 75*, 732-737.
- Williams, C. R., & Livingstone, L. P. (1994). Another look at the relationship between performance and voluntary turnover. *Academy of Management Journal, 37*, 269-298.
- Wright, R. E. (2000). Survival analysis. In L. G. Grimm & P. R. Yarnold's (Eds.) *Reading and understanding more multivariate statistics*. Washington, DC: American Psychological Association.

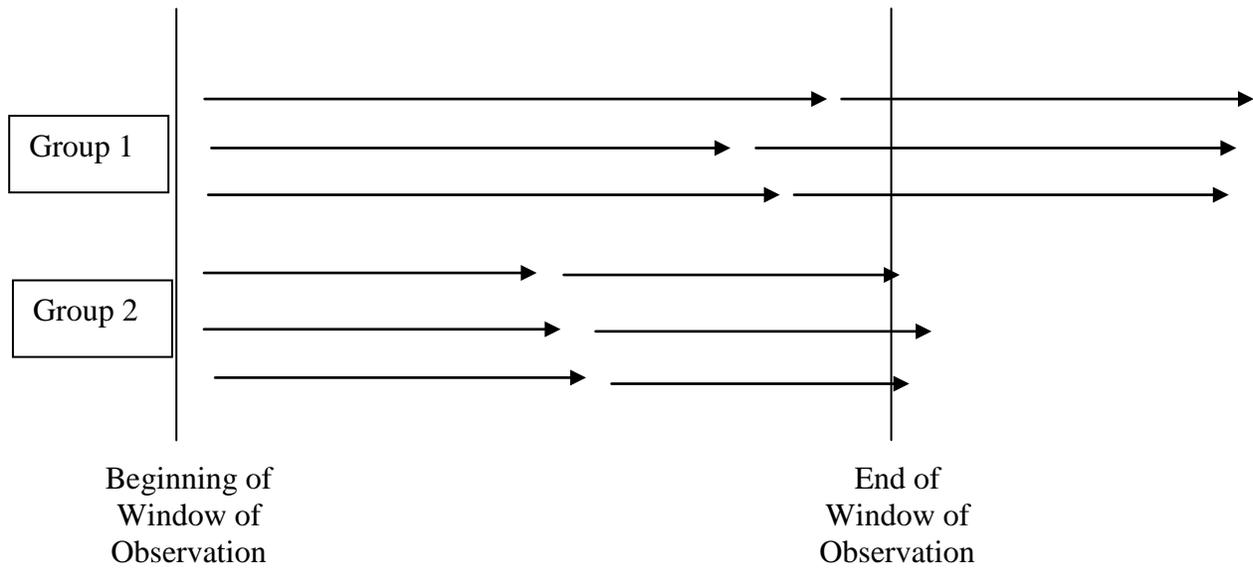


Figure 1. Two groups having the same turnover rate but different tenure distributions within group, leading to potential differences between Cox regression and linear and logistic regression. Length of each line represents tenure of an employee. Beginning of line represents hiring of an employee. End of line represents turnover. Vertical lines represent window of observation.

Table 1

Means, standard deviations, and correlations of the study variables

Variable ^a	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Sex	0.41	0.49	1												
2. White & Black vs. Others	0.22	0.16	-.035	1											
3. White vs. Black	0.20	0.44	.088	.097	1										
4. Age	31.20	9.58	-.194	-.059	.105	1									
5. Got preferred shift	0.82	0.39	.009	.035	.028	.099	1								
6. Has friends at company.	0.20	0.46	.071	-.002	-.010	-.129	.032	1							
7. Wage rate	9.28	0.50	.183	.011	.021	-.028	.090	.074	1						
8. Wage difference	1.25	2.27	-.185	.014	-.015	-.248	-.062	-.029	-.046	1					
9. Good history	0.00	1.73	-.032	.020	.009	.246	.024	.002	.080	-.252	1				
10. Dexterity test time	71.07	4.26	.254	-.001	-.082	.163	.001	-.015	.066	-.056	.022	1			
11. Appearance rating	0.00	0.45	.115	.001	-.091	-.059	-.019	.045	.048	-.112	.054	-.080	1		
12. Turnover	0.38	0.49	-.091	.085	-.036	-.003	.015	-.151	-.081	.007	-.069	-.127	-.070	1	
13. Tenure	313.17	224.41	.049	-.115	.007	-.010	-.082	.082	-.244	.020	.044	.012	.062	-.642	1

Note. Correlations are between nonmissing values of variables with missing values. (*N* ranged from 887 to 890)

Dichotomous/contrast variables have positively coded group named first or implied in name.

All correlations of .066 and higher are significant at $p < .05$

Table 2

Single-variable and multiple-variable Pearson r 's, logistic regression and Cox regression coefficients and associated p -values

Variable	Single-variable Analyses			Multiple-variable Analyses		
	Linear	Logistic	Cox	Linear	Logistic	Cox
	r^a	B^b	B^b	r^a	B^b	B^b
Sex	-.043	-.222	-.192	-.024	-.157	-.125
White & Black vs. Others	.077*	1.362*	1.289*	.085*	1.538*	1.306*
White vs. Black	-.003	-.016	.006	-.012	-.062	-.014
Age	.037	.009	.008	.037	.010	.006
Got preferred shift	.002	.014	.094	.001	-.005	.097
Has friends employed at company	-.135**	-.646**	-.553**	-.114**	-.569**	-.505**
Wage rate	-.071*	-.498*	-.022	-.058^	-.446^	.041
Wage difference	-.024	-.024	-.024	-.049	-.059	-.054^
Good employment history	-.065*	-.089*	-.062^	-.079*	-.118*	-.075*
Time to complete dexterity test	-.084*	-.057*	-.041*	-.100**	-.071**	-.050**
Appearance rating re. Interviewer mean	-.090**	-.461**	-.379**	-.094**	-.506**	-.383**

Note. All entries were computed controlling for type of job and for missing value indicators of those variables with missing values. Dichotomous/contrast variables have positively coded group named first or implied in name.

^a Semi-partial (part) r 's – the correlation of turnover with unique aspect of independent variable.

^b Regression coefficients.

^ $p < .10$. * $p < .05$. ** $p < .01$