

**Judgment and Choice
in Personnel Selection**

Working Paper #89-23

**Joel Rudin
Niagara University**

**John W. Boudreau
Cornell University**

Center for Advanced Human Resource Studies
Cornell University
393 Ives Hall
Ithaca, New York 14851-0952
(607) 255-2273

November, 1989

Comments welcome, please do not quote without permission from the author.

This research was carried out with support from the U.S. Army Research Institute, contract SFRC #MDA903-87-K-0001. The views, opinions and/or findings contained in this chapter are those of the author and should not be construed as an Official Department of the Army policy, or decision.

This paper has not undergone formal review or approval of the faculty of the ILR School. It is intended to make the results of Center research, conferences and projects available to others interested in human resource management in preliminary form to encourage discussion and suggestions.

JUDGMENT AND CHOICE IN PERSONNEL SELECTION

Behavioral decision theorists distinguish between processes of judgment (rating a set of alternatives) and choice (choosing the alternative with the highest rating). Studies have shown that subjects engaged in choice tasks reject all but a few alternatives on the basis of limited information. A re-analysis of previous research in personnel selection suggests an excellent ability of selectors in the choice task to identify the best candidate, whatever judgmental errors and biases they may display.

A simulated selection task carried out by 157 experienced managers confirms the excellent ability of subjects to choose the top candidate, and shows that subjects engaged in the choice task acquire less information about inferior candidates than do subjects engaged in the judgment task. Suggestions for future research are discussed.

Imagine that you have set out to buy a used car. You examine eight cars before making your choice, test driving some of them and rejecting others at first glance (due for example to excessive rust). A researcher asks you to rate each of the eight cars in terms of overall quality.

The researcher proceeds to sharply criticize you for carrying out an unsystematic search process. Your failure to test-drive every car and to ask the same questions to the dealers about each car has caused you to do a poor job of rank-ordering the cars. You respond that, since you could only afford one car, you had no interest in rank-ordering or in assigning ratings to the entire set of cars. It seems unfair to be criticized for poor performance of a task which was unrelated to your original mission of buying the best used car available.

This paper explores the possibility that a similar misspecification of the goals of employee selection has caused researchers to criticize selectors for behavior which may not adversely affect the goal of hiring the best individual from among a group of candidates.

Previous studies have not been designed with the goal of assessing subjects' ability to choose the best candidate from among a group. However, we can obtain some information from four studies which examine subjects' ability to rate hypothetical job candidates whose characteristics are experimentally manipulated (Bolster and Springbett, 1961; Hakel, Dobmeyer, and Dunnette, 1970; Constantin, 1976; Schuh, 1978). In all four studies, subjects assigned the most favorable ratings to the "best" candidates.

Two other widely cited studies (Wexley, Yukl, Kovacs, and Sanders, 1972; Wexley, Sanders, and Yukl, 1973) demonstrate a significant "contrast effect" in the assignment of scores to candidates of average quality. However, when these candidates were overrated they were always the best candidates in the group, and when they were underrated they were always the worst candidates in the group. Thus, the contrast effect served only to reinforce a correct choice.

The above studies form the empirical base for the widely held conclusion that unstructured interviews are prone to errors and biases caused by contrast effects and oversusceptibility to negative information (Hakel, 1982; Dipboye and Macan, 1988). Though these studies may demonstrate that errors and biases effect the scores assigned to candidates, they also suggest an excellent ability of selectors to choose the single best candidate from a group. Thus, if choice of the best candidate is the selector's objective (as opposed to the objective of accurately rating all candidates) the criticism of selectors expressed in these studies may be somewhat unfair.

How Often is it Necessary to Rank-Order Candidates?

It is often necessary to rank-order candidates when there is a possibility that an organization's top choice will refuse a job offer. If this happens, the organization will want to quickly be able to identify the next-best candidate in order to extend a job offer to him or her. It is also often necessary to rank-order candidates when many vacancies are to be filled simultaneously. If this is the case, the selectors will want to extend job offers in a "top-down" fashion (i.e. by first offering jobs to the best candidates). Indirect evidence suggests that many staffing decisions do not reflect these conditions.

Top candidates will usually accept job offers. Research in the area of job search and choice (summarized in Schwab, Rynes, and Aldag, 1987) suggests that only graduating college students in high demand fields can expect to have multiple job offers from which to choose. Even unemployed professionals tend to have to choose between a particular job offer and continued unemployment. In the internal staffing context, although it is expected that employees will be more likely to refuse promotions than they have been in the past (Hall, 1986), the desire for advancement is still strong among managers (Campbell and Moses, 1986). Promotion opportunities that do not disrupt existing work-family arrangements of employees, i.e. promotions that are not accompanied by geographic relocations, should still be highly desired by job candidates, and individuals who are offered such opportunities are likely to accept the offers.

Few selection decisions meet the condition of large numbers to be hired simultaneously. Most new jobs are created in organizations with less than twenty employees (Greene, 1982), and as organizations grow, their tendency is to create new job titles at a faster rate than the increase in new tasks to be performed (Baron and Bielby, 1986), which means that individual job titles in larger organizations will tend to be shared by few people.

Certainly, some important selection decisions, such as the hiring of letter carriers by the United States Postal Service (Adie, 1977), clearly require that all candidates be rank-ordered, because job offers may be rejected and many people are to be hired at once. But for many selection situations, rank-ordering of candidates cannot be assumed to be a worthwhile goal of selectors.

Thus, experiments in which subjects are required to rank-order candidates may lead to inaccurate conclusions and recommendations about appropriate behavior in choice tasks. This study addresses this problem by directly contrasting the behavior of personnel

selectors engaged in judgment (rank-ordering) and choice tasks. Fortunately, a large body of research suggests guidelines for predicting human behavior in situations which require that the best candidate be chosen from among a group.

Behavior in Choice Tasks

Several studies have been carried out in the field of consumer behavior relating to patterns of information acquisition and evaluation of alternatives in choice tasks. One group of studies (summarized in Punj and Stewart, 1983) examines the effect of a number of alternatives on strategies of information acquisition. When more than three alternatives initially exist, these studies have concluded that information acquisition and evaluation of alternatives follows a two-stage process. Initially, some alternatives are eliminated based on low levels of a small number of attributes. The remaining alternatives are compared in a compensatory manner, based on more information than was gathered on the initially rejected alternatives.

Johnson and Russo (1981) directly compared information acquisition patterns in judgment (rating) versus choice tasks, for student subjects who were instructed either to choose one of eight automobiles or to rate ("judge") each of the eight automobiles. They found that "the consumers who made choices used phased rules which eliminate alternatives, while consumers making judgments did not. (p. 154)"

Inferior alternatives are identified and rejected in the first phase of the choice process in accordance with the concept of "domination of alternatives". Alternative A is said to dominate Alternative B if the scores of A on all relevant dimensions are more favorable or at least as favorable as the scores of B. Alternative A will be preferred regardless of the relative contribution of each dimension to overall utility. If, on the other hand, A has a more favorable score than B on at least one dimension and a less favorable score than B on at least one other dimension, a conflict exists between the goals of maximizing overall utility and maximizing scores along each dimension. A and B form a conflict set (unless an Alternative C exists which dominates both A and B).

Though a failure to extract the maximum possible amount of information available may lead to rating errors, it will affect the quality of choices only if the initial screening leads to the rejection of the best possible alternative. To examine the likelihood of this, a series of Monte Carlo simulations was conducted by Johnson and Payne (1985). First, they found that simple choice rules were very effective at avoiding dominated alternatives, though little better than chance at choosing the best alternative. For large choice sets, a two-stage decision rule in which all but three alternatives were quickly eliminated by simple strategies, with more intensive analysis of the remainder was found to perform well with a fraction of the effort involved in more complex strategies.

Despite such evidence, researchers often recommend consistent treatment of job candidates. According to Gatewood and Field (1987, p. 371), "the major benefit of this consistency in questioning is that it makes comparison among applicants much easier." Taylor and Sniezek (1984) sharply criticize selectors for inconsistent treatment of candidates. They identify two sources of inconsistency: inconsistency among different

recruiters hiring for similar positions in terms of their beliefs of what topics should be covered in the interview, and inconsistency among individual recruiters between the topics that they believed to be important and topics which were actually discussed in the interview. Taylor and Snizek concluded (p. 166) that they "cannot help but question the validity of interview content that receives little agreement across a group of recruiters themselves and is covered in a manner highly inconsistent with individual recruiters' own importance ratings."

The evidence cited above suggests that for many selection situations where choice, not judgment, is the goal we should actually expect and encourage inconsistent treatment of candidates by selectors, and should not criticize this inconsistency unless it is demonstrated that the best candidates are rejected as a result.

Explaining Inconsistent Treatment of Job Candidates

Recently, it has been argued that insights from the field of interactionist psychology (IP) could be useful in explaining interviewer behavior (Eder and Buckley, 1988). They point out that during the employment interview, continuous feedback is provided to both the interviewer and the interviewee, resulting in constant modifications of behavior by both parties. Evaluative judgments of candidates are based on applicant characteristics, the selector's perceived function of the interview, cognitive limitations of selectors, and "interview process dynamics" (what goes on during the interview).

A complete interactionist perspective would also acknowledge that, at some point during the interview, the interviewer could correctly decide to reject the applicant. Once the selector is convinced that an applicant should be rejected, the selector might not bother asking additional potentially relevant questions. If the selector successfully eliminates dominated job candidates using questions that are asked early in the interview, the subsequent "inconsistent" treatment of dominated and non-dominated candidates will not impair the overall goal of choosing the best candidate.

The optimal order of questions to ask applicants has been mathematically formulated (Grether and Wilde, 1983). Questions should be asked in ascending order of the ratio of information cost to probability that this information could lead to a candidate's rejection. This strategy minimizes inspection costs by rejecting unacceptable candidates as early and efficiently as possible.

In the interview, there is little cost difference associated with asking one question rather than another, assuming that applicants can respond to either question in a similar length of time. Thus, the optimal strategy to winnow out unsuitable applicants is to start by asking questions with the lowest base rates (highest percentage of applicants whose answers to the questions render them to be unsuitable for the job).

Taylor and Snizek (1984) found a low level of consistency in interviewers' perceived importance of discussion of various interview topics, and found "that recruiters' low agreement can be attributed largely to their own differences of opinion about topic importance, not to the type of job opening (p. 162)." However, they seem to have

considered only job requirements and not base rates of applicants to be useful determinants of questioning strategies.

This lack of agreement could reflect rational, optimal behavior, if it reflects the fact that each company recruits from a somewhat unique pool of job applicants, not a random sample of the labor force (Boudreau and Rynes, 1985).

Recruiters from companies that are highly desirable to job applicants might find that the dimension with the lowest base rate among applicants concerned technical ability, so would initially focus on issues of technical competence. Recruiters from companies that are relatively undesirable to job applicants might initially focus on issues of job attraction (i.e. by asking candidates how they felt about certain potentially undesirable aspects of the job). Thus, inter-organizational differences in questioning strategies for similar jobs are to be expected, and may be quite consistent with the goal of choosing the single best candidate.

A final explanation for inconsistency in information acquisition patterns comes from Bettman and Park (1980) and Johnson and Russo (1981). They have found that for judgment (rating) tasks, the most experienced subjects acquired and processed the most information. For choice tasks, moderately experienced subjects acquired and processed the most information.

Hypotheses

While previous studies strongly suggest an excellent ability of subjects to make choices, most reported only the average ratings given to each candidate, not the number of subjects that chose or assigned the highest rating to each candidate.

The present study allows us to determine the number of subjects that choose or assign the highest rating to optimal candidates. This provides an opportunity for subjects to display their ability to identify the best candidates from among a group.

- H₁: Subjects will generally exhibit a preference for optimal (nondominated) candidates.
- H₂: Intra-interviewer inconsistency (inconsistencies by individual interviewers in their treatment of candidates) will be more pronounced in choice tasks than in judgment tasks.
- H₃: The best (non-dominated) candidates will receive the most intensive examination by interviewers.
- H₄: Inter-interviewer inconsistency (inconsistencies between different interviewers in their treatment of candidates) can be explained by differences in perceived base rates of candidates, such that interviewers will ask more questions on dimensions that, in their opinion, have lower base rates.

- H₅: Inter-interviewer inconsistency can also be explained by differences in previous experience making selection decisions. For judgment tasks, highly experienced subjects should be most intensive in their examination of candidates. For choice tasks, moderately experienced subjects should be most intensive.

Method

Procedure

Subjects were given two pieces of information (one each for the dimensions of Computer Competence and Interpersonal Skills) on each of eight "candidates" for promotion to the position of Accounting Supervisor. Eight additional pieces of information (four per dimension) about each candidate were obscured by removable labels. The number of removable labels peeled measured the extent of information acquisition. Half of the subjects were told to choose one of the candidates, while the other half were instructed to assign scores to the entire set of candidates.

The position of Accounting Supervisor was used because it was felt that such a position would be sufficiently generic to be found in any large or medium-sized organization.

The promotion task was used, rather than an external staffing task, because candidates for promotions in managerial ranks are usually unaware that they are actively being considered for a promotion (Markham, Harlan, and Hackett, 1987). This removes the need to consider other possible functions of selection devices, such as establishing good corporate public relations (Avery and Campion, 1982), that can affect strategies of information acquisition in an external staffing context.

The dimensions of Computer Competence and Interpersonal Skills were used because they reflected the technical and social aspects of managing professionals. Subjects in the pre-test indicated that they did not need information about additional applicant characteristics in order to make their decisions.

Descriptors along the two dimensions were designed to create positive, moderate, and negative impressions of candidates. The Interpersonal Skill descriptors were adapted from the Akron Leadership Questionnaire (Lord, Foti, and de Vader, 1984), which provides "prototypical", "neutral", and "antiprototypical" trait descriptions of leaders. Results from the pretest indicated that subjects were able to correctly distinguish the three levels of each attribute implied by the descriptors. A list of descriptors can be found in Rudin (1989).

All possible combinations of high, medium, and low levels along each dimension (except the combination of high interpersonal skills and high technical competence) were used to create the eight candidates. The two candidates who were high on one dimension and medium on the other represented the conflict set, and an "optimal" choice would involve the selection of one of these two candidates.

To examine the presence of contrast effects, and to control for the potential novelty effect of the task leading to high rates of information acquisition for the first few candidates, three different orders of candidate numbering were used. However, because all candidate information was presented on two pages that were entirely in view throughout the task, this research design did not ensure that subjects acquired information in the order in which candidates were numbered. Thus, the contrast effect is not tested as rigorously as in other studies (e.g. Schuh, 1978) in which subjects were required to assign scores to each candidate prior to receiving information on subsequent candidates.

A final manipulation involved priming the subjects that the base rate of candidates along one dimension was lower than the other. This was accomplished by advising subjects to assume that this selection task was originally to be conducted by someone who had left the company after collecting one piece of information per candidate per dimension. The original selector's "final piece of advice before leaving was that in his past experience he has found that", while most candidates had acceptable levels on one dimension, fewer had acceptable levels on the other dimension. In fact, base rates were identical for all dimensions and all manipulations. In every case, three of the eight candidates were given negative (unacceptable) descriptors along each dimension.

Thus, there was a total of twelve experimental conditions: Two tasks (judgment versus choice) X two base rate manipulations X three sequences of candidate numbering. Upon completing the task, subjects were instructed to refrain from peeling additional labels or re-attaching labels that they had peeled.

Subjects in the pretest were skeptical of the base rate manipulation. They felt that it was unlikely that the base rate on the dimension of computer competence would be lower than the base rate for interpersonal skills. To control for the possibility that subjects would choose to ignore the base rate manipulation, subjects were asked at the end of the exercise about their personal beliefs regarding the base rates. Subjects were asked "In general, which [dimension] do you believe would [an acceptable level] be harder to find among candidates for the position of Accounting Supervisor?"

Relevant information has been shown to play a greater role in judgments than irrelevant information (Constantin, 1976). To control for subjects' perceptions of the relevance of each dimension, subjects were asked "In general, which [dimension] do you believe would be more important for good performance in the position of Accounting Supervisor?"

Subjects' prior experience in making similar decisions was hypothesized to help explain inter-subject inconsistency. Two questions were aimed at assessing prior experience. One question asked subjects "Approximately how many times in your life have you played a part in deciding which one of a group of people should be hired or promoted to fill a vacancy?" The other measure of experience asked subjects how many years of full-time work experience they had.

The survey was distributed to managers in three large firms that are corporate sponsors of the Center for Advanced Human Resource Studies, as well as to employees of

Cornell University. The target audience was a mixture of human resource professionals and line managers. A high proportion of line managers was desired in order to test for the effects of formal education in Personnel/Human Resources Management on information seeking. Surveys were accompanied by a cover letter, signed by an executive associated with the subjects' organizations.

Results

Of the 230 exercises that were distributed, 157 were returned, a response rate of 74.8%.

Table One indicates the mean responses of the sample to the questions requesting information on subjects' backgrounds. The average number of years of full-time work experience was 19. The range of experience in making selection decisions was very broad (from 0 to "thousands"), providing large variability in the sample for detecting systematic effects of experience on information acquisition.

Table One: Descriptive Data (N = 157)

Mean Number of years full-time work experience	19.03 (7.36)
Extent of formal education in Personnel/HRM	
None	6.4%
Employer-sponsored training courses	77.1%
College-level courses	45.2%
Graduate school courses	46.5%
Proportion perceiving interpersonal skills to be more important for good performance	94.9%
Proportion perceiving interpersonal skills to be harder to find among candidates	94.9%

Note: Number in parentheses represents standard deviation.

Table Two shows the differences between tasks in amount of information acquired (measured as number of labels peeled). For total amount of information acquired, as well as amount of information pertaining to each of the dimensions of interpersonal skills and computer competence, significantly less information was acquired in the choice task than in the judgment task ($p < .01$). Subjects in the choice task peeled an average of over eight fewer labels than subjects in the judgment task (30.2 labels versus 21.6 labels).

Table Two: Mean Number of Labels Peeled Versus Task

Dimension	Task	
	Judgment	Choice
Interpersonal Skills	17.2	12.6
Computer Competence	13.0	9.0
Total	30.2	21.6

Note: All between-task means were significantly different at $p < .01$, according to t-tests. Total number of labels available to peel was 48 (24 per dimension).

Table Three shows the means and standard deviations of scores assigned to candidates by the subjects who were instructed to perform the judgment (rating) task. Analyses of variance showed that the sequence of candidate presentation never significantly affected the scores assigned to candidates, ($F_{\max} = 1.98$ for candidate HM). Interestingly, the scores assigned to the candidate who was preferred by almost all subjects, and who had the highest mean score (i.e., medium computer competence and high interpersonal skills) have a lower standard deviation than any other scores.

Table Three: Scores Assigned to Candidates (N = 77)

Candidate	Mean (Standard Deviation)	
MH	88.9	(9.6)
HM	69.6	(20.1)
LH	58.2	(24.3)
MM	57.6	(21.8)
LM	34.9	(23.8)
ML	26.0	(19.1)
HL	24.7	(20.6)
LL	12.0	(15.6)

Note: All differences between means are statistically significant at $p < .05$, except for the following pairs of candidates: LH and MM, and ML and HL.

The first letter identifying the candidate reflects the level of computer competence (High, Medium, or Low). The second letter identifying the candidate reflects the level of interpersonal skills. Actual descriptors used to identify candidates are listed in Rudin (1989).

Table Four: Preference for Candidates by Task and Order (N = 154)

Choice Task				
Order	MH	Candidate HM	LH	MM
1 (26)	96	0	4	0
2 (29)	100	0	0	0
3 (25)	92	4	0	4

Judgment Task				
Order	MH	Candidate HM	LH	MM
1 (26)	88	12	0	0
2 (25)	88	12	0	0
3 (23)	96	4	0	0

Note: The first letter identifying the candidate reflects the level of computer competence (High, Medium, or Low). The second letter identifying the candidate reflects the level of interpersonal skills. Actual descriptors used to identify candidates are listed in Rudin (1989).

The "preferred" candidate in the judgment task was the candidate who was assigned the highest score. Table Two indicates the three orders of candidate presentation that were used.

Number in parentheses in first column represents row frequencies (number of subjects in each task/order combination). Cell entries represent percentage of subjects in each task/order combination who preferred a particular candidate.

Table Four shows the distribution of preferred candidates by task and order. In the choice task, a candidate was considered to be preferred if the candidate was chosen by the subject. In the judgment task, a candidate was considered to be preferred if s/he had the highest score.

Table Four shows that 152 of the 154 subjects who indicated a preference chose members of the dominant set of candidates (the two candidates who could possibly represent the optimal choice). 144 subjects preferred the candidate with high interpersonal skills and medium computer competence, while 8 subjects preferred the candidate with high computer competence and medium interpersonal skills. Preferences were clearly impervious to the sequence of candidate presentation and to the task. Thus, H_1 was very strongly confirmed.

Candidate quality was highly related to the amount of information acquired on each candidate. Table Five shows the differences in mean amounts of information acquired (measured as number of labels peeled) along each dimension for each task. For the dimension of interpersonal skills, a direct relationship between applicant quality and amount of information acquired was found, with a large and significant decline in information acquired as quality falls from medium to low. Analyses of variance indicated a significant interaction effect between task and candidate quality on information acquisition patterns ($p < .05$). The effect of applicant quality was stronger for the choice task than the judgment task, confirming H_2 for this dimension.

Table Five: Information Acquired Versus Applicant Quality,
by Task

Choice Task (N = 80)		
Dimension		
Quality	Computer Competence	Interpersonal Skills
High	2.15	1.00
Medium	1.88	1.53
Low	0.90	0.81

Judgment Task (N = 77)		
Dimension		
Quality	Computer Competence	Interpersonal Skills
High	2.52	1.40
Medium	2.39	1.75
Low	1.68	1.65

Note: Information acquisition was measured as number of labels peeled (maximum of three labels per candidate on each dimension). Between-task differences in information acquisition are displayed graphically in Figures One (Interpersonal Skills) and Two (Computer Competence).

Within each task, between-quality differences in information acquisition are statistically significant at $p < .05$ except in the following cases: Judgment task, Interpersonal Skills, High versus Medium quality; Judgment task, Computer Competence, Medium versus Low Quality.

For the dimension of computer competence, the effect of applicant quality is again stronger for the choice task than the judgment task, but Table Five demonstrates that the relationship is nonmonotonic, as applicants who are high on this dimension were the targets of significantly less information acquisition than candidates who were medium on this dimension. Analyses of variance indicated a significant interaction effect between task and quality on information acquisition patterns ($p < .01$). The effect of applicant quality was again stronger for the choice task than the judgment task, confirming H_2 for this dimension as well.

Hypothesis H_3 predicted that the members of the conflict set would be the targets of more intensive search than other candidates, particularly in the choice task. Table Six shows the information acquisition on each candidate and task. Generally, the most highly desired candidate, candidate MH, was the target of the most information acquisition, and had similar amounts of information acquired by subjects engaged in both tasks (approximately 4.9 labels peeled). Five of six dominated candidates have significantly less information acquired about them in the choice task than in the judgment task, confirming hypothesis H_3 .

Inter-interviewer differences had no significant effect on information acquisition patterns. For each task, a regression was carried out to determine the association of amount of information acquired along each dimension with the independent variables of work experience, experience making selection decisions, the base rate manipulation, and extent of formal education in Personnel/Human Resources. No model approached statistical significance ($F_{\max} = 1.48$), as shown by Table Seven, so hypotheses H_4 and H_5 were not confirmed.

Table Six: Levels of Information Acquisition by Candidate and Task (N = 157)

Candidate	Task	
	Judgment	Choice
MH	4.98	4.88
LH*	4.61	2.69
HM*	4.55	3.95
MM	4.03	3.75
LM*	3.58	1.56
ML*	3.22	1.80
HL*	2.78	1.69
LL*	2.58	1.25

* - between-task differences are significant at $p < .05$

Note: Information acquisition was measured as number of labels peeled (maximum of six per candidate).

The first letter identifying the candidate reflects the level of computer competence (high, medium, or low). The second letter identifying the candidate reflects the level of interpersonal skills. Actual descriptors used to identify candidates are listed in Rudin (1989).

Table Seven: Inter-Individual Differences in Information Acquisition

	Judgment (N = 77)		Choice (N = 80)	
	Interpersonal Skills	Computer Competence	Interpersonal Skills	Computer Competence
F	1.48	1.23	1.40	0.88
Adj. R ²	0.04	0.02	0.03	-0.01
Selection Experience	-1.05	-1.23	-0.15	-0.87
Base Rate Prime	1.50	2.50	-4.20*	-2.11
No PHR Education	-3.71	-2.57	-0.20	-0.01
College PHR Education	1.67	1.43	-0.99	-1.52
Graduate PHR Education	-2.02	0.66	-0.73	-0.31
Work Experience	0.08	0.15	0.05	-0.03

* - significant at $p < .05$

- Dependent variables are number of labels peeled along each dimension (maximum of 24). Cell entries represent regression coefficients associated with each independent variable. Base Rate prime was a dummy variable, set to 1 when subjects were advised that fewer candidates had acceptable computer competence than suitable interpersonal skills.

Discussion

This experiment has demonstrated similarities and differences between judgment and choice in personnel selection. The processes differ in that the choice task is carried out using significantly less information than the judgment task, but they converge in that the same candidate is preferred by subjects engaged in both tasks, and the same amount of information is acquired on the preferred candidate by subjects engaged in both tasks.

The choice task can be thought of as a simplified version of the judgment task. Whereas the goal of the choice task is simply to identify the best available candidate, the judgment task adds the goal of correctly rank-ordering the remaining candidates. The higher overall rate of information acquisition in the judgment task is caused by the more thorough examination of dominated candidates.

This study corroborated previous evidence about the quality of selector preferences. The same candidate was preferred by the overwhelming majority of subjects, and nearly every subject preferred a member of the dominant set of candidates. In conjunction with the previous studies in which applicant quality has been experimentally manipulated and subjects have correctly identified the best candidate, a total of over 500 subjects have been tested, and this finding consistently emerges.

Further evidence that subjects did a better job of identifying the best candidate than rank-ordering the candidates comes from the fact that the scores assigned to Candidate MH, who was preferred by almost all subjects, exhibited less variation than the scores assigned to other candidates. Recently, it has been argued that the effectiveness of unstructured interviews has been underestimated by research designs which pool results across subjects (Dreher, Ash, and Hancock, 1988). Results of this study suggest that interview effectiveness may also be underestimated by pooling results across both dominated and non-dominated job candidates, because the scores assigned to Candidate MH were more reliable than the other scores, and high reliability is a necessary though not sufficient condition for high validity (Guion, 1965, p. 31).

Whatever judgmental errors and biases they may display, subjects appear to be able to correctly identify the best candidate from among a group of candidates. Thus, for selection decisions that do not require rank-ordering of candidates, the results of this study combined with a re-interpretation of the results from the previous five studies strongly suggest that selectors perform surprisingly well, when faced with choice tasks.

Results of this study also strongly support a revision of the perspective of interactionist psychology as it applies to personnel selection, to incorporate the notions that choice (as opposed to judgment) is a possible goal of selection and that the decision to reject the candidate can be made during the course of the interview. In choice tasks, it is incorrect to assume that subjects erred in some way by acquiring little information on dominated candidates, because evidence shows that the quality of the choices by such subjects has been uniformly high. Likewise, it is incorrect to unequivocally criticize selectors for inconsistent treatment of candidates in situations for which candidates need not be rank-ordered, if this inconsistency is caused by differences in applicant quality.

Many authors have recommended structured interviews (cites), but organizations which train their selectors to use structured interview formats should first determine whether the selectors need to rank-order job candidates. If not, organizations may more profitably devote their training dollars toward goals other than accurate ratings of candidates who stand no chance of being hired.

Inter-interviewer inconsistencies were not associated with the independent variables used in this study, but an experimental manipulation of base rates will only be successful if the manipulation is perceived to be credible by subjects (Scholz, 1987). Results strongly suggest that subjects in this study did not use the base rate information because they did not believe it. Nearly all subjects felt that interpersonal skills were not only more important for good performance but also harder to find among candidates. The fact that actual experimental base rates were equal may also have contributed to this effect.

The greater amount of information acquired by subjects in both tasks on the dimension of interpersonal skills may indicate that, although subjects ignored the experimental condition, their behavior was "rational" in view of their personal perceptions of base rates. On the other hand, subjects may have acquired more information about interpersonal skills because they felt that additional information pertaining to computer competence would not assist them in evaluating candidates. A true test of the effect of base rates on information acquisition patterns will be possible only in an environment where there is greater diversity of opinion between subjects' base rate perceptions.

The reduction in information acquisition for candidates with high computer competence suggests that subjects were screening along this dimension. A candidate who initially appeared to have high computer competence would pass this screening phase, while additional information would be required for other candidates to determine whether or not their levels of computer competence were acceptable.

The dimension of interpersonal skills appeared to play a greater role in the rank-ordering of candidates, since the three candidates who received the lowest ratings were the three candidates with low interpersonal skills, and the two candidates with high interpersonal skills were ranked first and third. For this dimension, low interpersonal skills apparently served as a signal to selectors that the candidate was dominated by others, so an initial indication of low interpersonal skills led to a decrease in additional search.

Thus, the relationship between negative information and subject behavior in this study appears to be as follows: For the dimension of prime interest to the selector, negative information reduces information acquisition. For other dimensions, positive information reduces information acquisition.

This result does not necessarily conflict with previous research which found that negative information reduces information acquisition, because previous studies (e.g. Bolster and Springbett, 1958) did not allow subjects to search independently along different dimensions, so it was impossible with previous research designs to determine interdimensional differences on the effect of quality on information acquisition.

This study failed to find evidence of an effect of sequence of candidate presentation on either the assignment of scores to candidates or preference for one candidate. Although this study does not test for contrast effects as stringently as possible, because subjects

acquire their information on candidates prior to assigning scores to the candidates, no study has ever demonstrated contrast effects for more than four candidates (Landy and Bates, 1973), so it is possible that the contrast effect is relevant only in situations characterized by four or less candidates and a need to rank-order the set of candidates. While these situations can occur, organizations should ensure that their selectors face such situations before spending money on the intensive training needed to alleviate this judgmental bias (Wexley, Sanders, and Yukl, 1973).

Suggestions for Future Research

This study demonstrates the importance of adopting a more flexible view of the nature and purpose of employee selection (Cronbach and Gleser, 1965), to avoid criticizing selectors for behavior that is in fact irrelevant to their tasks. Employee selection situations can be categorized as follows:

1) Pure Judgment. This is a situation in which all candidates must be rank-ordered. For this class of selection situations, the validity coefficient and related utility measures are appropriate measures of effectiveness.

2) Pure choice. In this situation, there is no need to rank-order candidates, because there is only one vacancy and little chance that it will be refused by the preferred candidate. Many promotion decisions fall into this category, as well as many external staffing decisions for smaller organizations. All available evidence suggests that this task is performed exceptionally well by subjects. Errors and biases that are associated with the task of rating inferior candidates may depress validity coefficients without adversely affecting the quality of decisions made in this situation. Utility estimates based on differences in validity may overestimate gains from high-validity procedures.

3) Mixed judgment/choice. In this situation, it is necessary to rank-order a subset of the available candidates (i.e. the top three candidates). Previous research offers little insight as to subjects' ability to identify the second-best or third-best candidates. Previous studies have used a conflict set of one candidate, making it impossible to determine the "correct" rank-ordering of other candidates. However, in this study, there was a conflict set of two candidates, and these two candidates received the two highest scores. This suggests that subjects may be able to effectively "sort out the wheat from the chaff" and identify the best candidates, without accurately rating all candidates.

Thus, there is a tremendous need for contingency-based research in the area of personnel selection, so that prescriptive advice that is provided by personnel psychologists fits the task environment. It is no longer sufficient to advocate consistent treatment of all candidates at all times, through selection devices such as structured interviews. Consistency is necessary and efficient only when rank-ordering of candidates is necessary.

The contingent nature of employee selection also has implications for the assessment of discrimination in employment. Currently accepted measures of discrimination assess the degree to which candidates of protected and non-protected groups are judged fairly and consistently (Arvey and Faley, 1988). In situations which do not necessitate the rank-ordering of candidates, different measures are needed, such as comparisons of the frequency of choice of protected versus non-protected group members, (e.g., "adverse impact").

The present study was unsuccessful in establishing sources of inter-interviewer inconsistency. Perceived base rates of candidates faced by different interviewers seem to be a potentially useful explanatory variable, as indicated in this study by the more intense search for information on the dimension of interpersonal skills, but the present study's base rate manipulations were not perceived to be credible by subjects.

The lack of responsiveness of subjects to our base rate manipulation raises the possibility that selectors may be overly resistant to changing their perceptions of base rates as environment change. This could have potentially serious effects on selection effectiveness, as the number of recommendations of unsatisfactory candidates may increase as perceptions of base rates become further removed from reality. How do selectors form hypotheses as to the base rates of candidates on various dimensions, how accurate are these hypotheses, and what does it take to change these perceptions? The answers to these questions are crucial to organizations that are recruiting in novel or turbulent environments (i.e. a multinational corporation establishing a facility in a country in which it has never done business before, or facing demographic changes in its applicant pool).

The results of this study suggest that there has been an overemphasis on negative outcomes in previous research in the area of personnel selection. However, this overemphasis has been on the part of researchers, not subjects. In this study, and previous studies that manipulated applicant qualifications experimentally, subjects correctly identified the best candidates. Yet, previous research has dwelled on biases in scores assigned to candidates of moderate and low quality. Available evidence suggests that the task of choosing the best candidate from among a group is one that is performed well.

References

- Adie, D. (1977). An Evaluation of Postal Service Wage Rates. Washington: American Enterprise Institute.
- Arvey, R. and Campion, J. (1982). The Employment Interview: A Summary and Review of Recent Research. Personnel Psychology, 35:281-322.
- Arvey, R. and Faley, R. (1988). Fairness in Selecting Employees. Reading, MA: Addison-Wesley.
- Baron, J. and Bielby, W. (1986). The Proliferation of Job Titles in Organizations. Administrative Science Quarterly, 31:561-586.
- Bettman, J. and Park, C. (1980). Effects of Prior Knowledge and Experience and Phase of the Choice Process on Consumer Decision Processes: A Protocol Analysis. Journal of Consumer Research, 7:234-48.
- Bolster, B. and Springbett, R. (1961). The Reaction of Interviewers to Favorable and Unfavorable Information. Journal of Applied Psychology, 45:97-103.
- Boudreau, J. and Rynes, S. (1985). Role of Recruitment in Staffing Utility Analysis. Journal of Applied Psychology, 70:354-66.
- Campbell, R. and Moses, J. (1986). Careers from an Organizational Perspective, in Hall, D. (Ed.) Career Development in Organizations. San Francisco: Jossey-Bass.
- Constantin, S. (1976). The Investigation of Information Favorability in the Employment Interview. Journal of Applied Psychology, 61:743-9.
- Cronbach, L. and Gleser, G. (1965). Psychological Tests and Personnel Decisions. Urbana: University of Illinois Press.
- Dipboye, R. and Macan, T. (1988). A Process View of the Selection/Recruitment Interview. In Schuler, R., Youngblood, S., and Huber, V. (Eds.) Readings in Personnel and Human Resource Management. St. Paul: West.
- Dreher, G., Ash, R., and Hancock, P. (1988). The Role of the Traditional Research Design in Underestimating the Validity of the Employment Interview. Personnel Psychology, 41:315-325.
- Eder, R. and Buckley, M. (1988). The Employment Interview: An Interactionist Perspective, in Rowland, K. and Ferris, G. (Eds.), Research in Personnel and Human Resources Management, vol. 6. Greenwich, CT: JAI.
- Gatewood, R. and Field, H. (1987). Human Resource Selection. New York: CBS.
- Greene, R. (1982). Tracking Job Growth in Private Industry. Monthly Labor Review, 105:3-9.

- Grether, D. and Wilde, L. (1984). An Analysis of Conjunctive Choice: Theory and Experiments. Journal of Consumer Research, 10:373-85.
- Guion, R. (1965). Personnel Testing. New York: McGraw-Hill.
- Hakel, M. (1982). Employment Interviewing. In Rowland, K., and Ferris, G. (Eds.) Personnel Management. Boston: Allyn and Bacon.
- Hakel, M., Dobmeyer, T., and Dunnette, M. (1970). Relative Importance of Three Content Dimensions in Overall Suitability Ratings of Job Applicants' Resumes. Journal of Applied Psychology, 54:65-71.
- Hall, D. (1986). Introduction: An Overview of Career Development Theory, Research, and Practice, in Hall, D. (Ed.) Career Development in Organizations. San Francisco: Jossey-Bass.
- Johnson, E. and Payne, J. (1985). Effort and Accuracy in Choice. Management Science, 31:395-414.
- Johnson, E. and Russo, J. (1981). Product Familiarity and Learning New Information. Advances in Consumer Research, 8:151-5.
- Landy, F. and Bates, F. (1973). Another Look at Contrast Effects in the Employment Interview. Journal of Applied Psychology, 58:141-4.
- Lord, R., Foti, R., and De Vader, C. (1984). A Test of Leadership Categorization Theory: Internal Structure, Information Processing, and Leadership Perceptions. Organizational Behavior and Human Performance, 34:343-78.
- Markham, W., Harlan, S., and Hackett, E. (1987). Promotion Opportunity in Organizations: Causes and Consequences, in Rowland, K. and Ferris, G. (Eds.) Research in Personnel and Human Resources Management, Vol. 5. Greenwich, CT: JAI.
- Punj, G. and Stewart, D. (1983). An Interaction Framework of Consumer Decision Making. Journal of Consumer Research, 10:181-96.
- Rudin, J. (1989). Judgment and Choice in Personnel Selection. Cornell University: Unpublished Ph.D. Dissertation.
- Scholz, R. (1987). Cognitive Strategies in Stochastic Thinking. Dordrecht: Reidel.
- Schuh, A. (1978). Contrast Effect in the Interview. Bulletin of the Psychonomic Society, 11:195-6.
- Schwab, D., Rynes, S., and Aldag, R. (1987). Theories and Research on Job Search and Choice, in Rowland, K. and Ferris, G. (Eds.), Research in Personnel and Human Resources Management, Vol. 5. Greenwich, CT: JAI.
- Taylor, M. and Sniezek, J. (1984). The College Recruitment Interview: Topical Content and Applicant Reactions. Journal of Occupational Psychology, 57:157-168.

Wexley, K., Yukl, G., Kovacs, S., and Sanders, R. (1972). Importance of Contrast Effects in Employment Interviews. Journal of Applied Psychology, 56:45-8.

Wexley, K., Sanders, R., and Yukl, G. (1973). Training Interviewers to Eliminate Contrast Effects in Employment Interviews. Journal of Applied Psychology, 57:233-7.