DOCUMENT RESUME

ED 396 386 EA 027 610

AUTHOR Johanson, George A.; Gips, Crystal J.
TITLE Secondary School Hiring Preferences: An Exploratory Study.
PUB DATE Apr 90
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Administrator Attitudes; *Cognitive Processes; *Interpersonal Competence; Principals; Secondary Education; *Secondary School Teachers; Social Cognition; Teacher Competencies; Teacher Employment; *Teacher Selection

ABSTRACT

This paper presents findings of a national survey that asked principals to identify qualities they found desirable in teaching candidates. A total of 271 out of 500 secondary principals returned completed questionnaires, a 54 percent response rate. A single questionnaire item, in which respondents rated a fictitious teaching candidate, elicited a 77 percent response rate. Responses were measured by three kinds of formats--Likert, forced-choice, and free-response. The analysis developed a composite scaling of teaching candidate qualities that correlated highly with each format. In general, the principals greatly valued affective qualities over cognitive qualities. In selecting the most capable candidate, principals preferred candidates with greater cognitive capabilities when affective qualities were held constant. However, principals viewed the cognitive qualities as relatively unimportant compared with the affective qualities--so much that cognitive attributes might well prove extraneous in practice. Finally, the way in which a question was asked may have a large effect on both responses and the possible interpretations of those responses. Five tables are included. (Contains 18 references.) (LMI)

SECONDARY SCHOOL HIRING PREFERENCES: AN EXPLORATORY STUDY

George A. Johanson & Crystal J. Gips
College of Education, Ohio University
Athens, Ohio

A paper presented at the annual meeting of the American Educational Research Association
Boston, MA
April 1990

Abstract

What are the qualities that secondary school principals desire in a teaching candidate? Is it possible to accurately assess the relative importance of these qualities independently of the format of the questions? Do principals truly desire the 'best and brightest' candidates? A national survey of principals identified nine qualities that principals perceived as most desirable in teaching candidates and preferences towards these qualities were measured using three item formats: Likert, forced-choice, and free-response. A composite scaling of these qualities was derived that was highly correlated with each format. In general, the principals greatly valued affective qualities over cognitive qualities. The question of selecting the most capable candidate proved to be somewhat more complex. While there was a definite preference for candidates with greater cognitive capabilities when affective qualities were held constant, the cognitive qualities were seen as so relatively unimportant compared to the affective qualities that they might well prove extraneous in practice. Implications for educational methodology, practice and reform are discussed.

SECONDARY SCHOOL HIRING PREFERENCES: AN EXPLORATORY STUDY

The most capable young persons do not enter public school teaching in large numbers. Furthermore, of those who do begin careers in education, the most able are more likely leave (Provalko, 1970; Schlecty & Vance, 1981; Weaver, 1979). Recent national studies in education by various individuals and groups (e.g., the Holmes Group) urge that we make a stronger effort to attract and retain the 'best and brightest' individuals. Indeed, current research into teaching and learning confirms what many have suspected for a long time: the most effective teachers thoroughly understand the subject area themselves. Suggested educational reforms frequently include a call for increased efforts in both recruitment to, and retention in, teacher education.

Standing between those newly recruited to teacher education and a long successful career in the classroom, however, is the selection and hiring process. Several studies of the selection of new teachers give evidence that hiring preferences may be subverting efforts to attract the most proficient. In particular, David Berliner (1987) responded to comments by principals that they do not always seek the brightest individuals:

> It is a rare field where a practitioner would be willing to say "the smarter you are the worse you'll do." Yet in education this belief has vocal supporters. Principals with such beliefs hold the most tenuous logic and have no research evidence with which to defend their beliefs. To an outsider, it sounds bizarre and represents clear discrimination against the hiring of the brightest teachers. (p. 22)

Recent research goes beyond beliefs to demonstrate that principals do, in fact, act on the convictions they express. In a comparison of National Teacher Examination (NTE) scores and supervising teacher ratings of recent student teachers who sought employment as teachers, Browne and Rankin (1986) found a significant positive correlation between NTE scores and the student teacher ratings, but did not find a significant correlation between NTE scores and employment. Perhaps most interesting was the finding of a negative relationship between the rating of the student teachers on the characteristic termed "intellectual brightness" and the likelihood that they would have obtained a job as a teacher. Are principals really biased against the most capable candidates for teaching positions in their schools, or is this just an unfortunate impression confounded by a misunderstood belief structure?

Some researchers have attempted to identify teacher characteristics and beliefs which lead to high levels of student learning through assessment of teacher behaviors in the classroom (Brophy, 1982; Leinhardt & Greeno, 1986; Sandefur & Adams, 1976). Others have tried to assess those factors principals and others seem to value in teachers (Gips & Bredeson, 1985; Henry & Sa'ad, 1977; Johnson & Prom-Jackson, 1986; Kowalski & Weaver, 1989). In general, research findings indicated that intellectual skills were

less valued than were affective factors, but the factors or qualities examined tended to vary from study to study as did the item format and scaling approach.

Method

From the research literature, a survey instrument was devised and piloted (Gips & Johanson, 1989) in an effort to both identify the characteristics deemed most important to principals in the hiring of a teacher and to rate the similarity of those characteristics to each another. The preliminary list included four affective traits: concern & caring (C&C), rapport (RAP), enthusiasm (ENT), communication/instructional skills (CIS) and two cognitive traits: subject-area knowledge (SAK) and intellectual capacity (INT). The pilot study (using 60 current secondary principals) had a free-response item that asked 'What is the single most important quality you look for in a candidate for a teaching position at your school?' From this item and respondent's comments, three additional traits were identified: dedication to the profession (DED), a cooperative & flexible attitude (C&F), and integrity & character (I&C).

Pairwise ratings of the similarity of the six original characteristics (a non-metric one-dimensional scaling yielded values ranging from 0.669 to 0.728 for the four affective traits and values of -1.407 and -1.420 for the two cognitive traits) verified that the affective/cognitive division suggested by previous studies was appropriate. That is, the affective characteristics were seen to be very similar to one another and very different from the cognitive characteristics which were also seen to be very similar to one another.

The Survey

For this study, the decision was made to use both a Likert scale and paired comparisons for rating the importance of the nine previously identified traits in addition to the same free-response question as in the pilot study.

The reasons for this approach were as follows: 1. the pilot had shown a 'ceiling effect' in the Likert scaling of the preferences for the affective characteristics since all were seen as very desirable 2. the pairwise comparisons with a Thurstone Case V scaling (Dunn-Rankin, 1983) would yield an interval level scale that would permit stronger interpretations 3. the forced-choice format of the pairs could be compared with the unforced Likert scaling 4. the free-response format would permit both a scaling of the traits that might more accurately reflect practice and would indicate if the nine previously identified characteristics formed a relatively complete list, and finally, 5. if principals had clearly defined priorities, the methods should converge to a reasonable extent.

The pilot had indicated that the trait identified as cooperative & flexible (C&F) may have been sometimes interpreted to

mean 'cooperation with the educational structure' and at other times interpreted as meaning 'cooperation with students'. To further explore this, the Likert item regarding C&F was split into two separate items.

A question regarding the desirability of a (fictitious) candidate was used with two levels of grade point average (GPA) and with an undergraduate degree from two universities with differing socioeconomic or institutional status to determine if there was, in fact, a preference for the less able candidate with respect to either GPA or institutional status. This necessitated four forms of the survey, but permitted random assignment of subjects to forms.

Finally, there was a series of questions regarding personal strengths of the respondents themselves, a series of questions alluding to possible reasons for preferring a less cognitively able candidate, some demographic items, and a summative item with a Likert format: 'It is critically important to attract the "best and brightest" persons to careers in teaching.'

The Sample

A simple random sample of 500 secondary principals from across the United States was sent the survey. A second mailing resulted in a total of 271 usable responses (54.2%). A third mailing to the 229 non-respondents asking only for the rating of the fictitious candidate resulted in an additional 117 (51.1%) usable responses to this item. That is, for this single item, the response rate was 388 out of 500 (77.6%).

Results

Demographic data on the respondents indicated that the sample fit common descriptors of secondary school principals with regard to mean age (46.3), years of experience as a principal (10.1), and school size (804.1). The gender distribution was representative of the male/female ratio among secondary school principals (90.7% male).

Principals felt that they had a major role in the selection of teachers in their schools. On a scale of 1 to 7 (1: 'the final decision is entirely out of my hands' to 7: 'the final decision is entirely mine') the mean was 5.75 (s.d. 1.2) and 50% of the principals were at the mode, 6. In particular, political factors in the selection were seen as unimportant. It seems fair to conclude that the principals in the sample perceived themselves as having considerable control over the selection of teachers. It should be noted that, especially in larger districts, the transfer of teachers from one school to another may be beyond the control of the principal.

Likert Scale

The ten declarative Likert statements had options: Strongly Agree (5), Agree (4), Neutral (3), Disagree (2), and Strongly

Disagree (1). The mean responses to the two items regarding cooperation were deemed sufficiently similar (a mean of 4.17, s.d. of 0.72 for 'cooperation with the educational structure'; a mean of 4.08, s.d. of 0.67 for 'cooperation with students') to combine and the mean of the two will be subsequently referred to as cooperation and flexibility (C&F).

Intellectual capacity (INT) was rated as least important and concern and caring (C&C) as most important. Table 1 shows the mean

Insert Table 1 about here

response and the standard deviation for each item. Even though all of the statements were more strongly worded than in the pilot study, for example, '--would tend to eliminate from further consideration any teaching candidate who failed to demonstrate a true dedication--', the mean response to the nine characteristics was 3.99 or general agreement. In particular, there were only 17 instances of 'Strongly Disagree' with the any of the ten items (0.63%) and 8 of these ratings came from just two respondents.

A summated score was computed from all ten Likert items for each respondent. The reliability of this summated scale was 0.79 (Cronbach's coefficient alpha). A principal component analysis yielded three factors with a total of 58.2% of the variation explained by these dimensions. From the rotated (Varimax) factor loadings, the first factor (35.8%) was clearly affective, the second (12.2%) primarily cognitive, and the third (10.3%) had only the two cooperative items loading on it. In short, there was some evidence of reliability and construct (factor) validity for the summated Likert scale.

A dichotomous scaling of the Likert responses was performed in which 'Strongly Agree' and 'Agree' were coded similarly and distinct from 'Neutral', 'Disagree', and 'Strongly Disagree' in an effort to detect both positional changes, if any, and the effect of using a five-point scale versus a (pseudo) two-point scale. This rescaling is labelled 'Agree' in Table 1 and the scale values represent the proportion of respondents who agreed with each statement. Note that all but three traits have agreement at or above the 90% level and that the two cognitive traits, SAK and INT are agreed with 65% and 41% of the time, respectively. There was one positional change; CIS was ranked fourth using the five-point scale, but ranked second using the two-point scale. As might be expected, overall agreement between the two-point and five-point scales was very high; Table 2 shows the correlation, 0.99.

Insert Table 2 about here

Paired Comparisons

The Thurstone scale values from the paired comparisons for the nine teacher characteristics are also seen in Table 1. Integrity and character (I&C) was seen as most important and dedication to the profession (DED) as least important. The zero value for dedication is arbitrary and only indicates relative position. Note that I&C is approximately twice as far from dedication as the highest rated cognitive measure, subject-area knowledge (SAK) and five times as far from dedication as intellectual capacity (INT). That is, relative to dedication, integrity and character was seen as twice as important as subject-area knowledge and five times as important as intellectual capacity. While C&C was rated most important in the Likert scaling and third in the forced-choice format, the distance of C&C from DED is still approximately four times the distance of INT from DED.

An often used measure of the strength of a linear (interval preserving) relationship is the (Pearson) correlation coefficient. Even with the altered rankings indicated above, the Thurstone scale values correlated 0.83 with the Likert scale values and 0.81 with the dichotomous Agree scale (Table 2).

When paired comparisons are made, it is always possible that the overall ordering of a particular pair of characteristics may differ from the ordering from the 'head-to-head' comparison as a pair. There was only one pairwise comparison in which the majority of the respondents disagreed with the overall ordering of the traits. The item was the comparison of concern and caring with communication/instructional skill. Overall, CIS was seen to be slightly more important that C&C, however, C&C was selected more often than CIS in direct comparison. To see if there was substantial agreement with the final ranking of the traits, each response was rescored as being 'correct' (in agreement with the overall ranking) or 'incorrect'. The mean score for all subjects over all pairs was 70%. That is, on average, a respondent had 25 of the 36 paired comparisons in the direction of the final ranking.

The Free-Response Question

While the free-response format neither forces choices nor provides qualities to be rated, it does present unique coding problems (Baldwin, et al., 1988). The responses to the item asking for 'the single most important quality' sought in a teaching candidate were classified separately by each of the authors. Differences were discussed until a consensus classification was reached. In all, 258 of the 271 respondents indicated one or more qualities. One hundred ninety-six of the 258 first-mentioned qualities were classified as one of the nine anticipated qualities. Table 1 contains the numbers and proportions of each of the nine. The remaining 62 responses were classified as: referring to credentials and/or previous experience (24), pejorative (19), too broad or general to have real meaning (7), and miscellaneous (12). The following examples will give a sense of the additional

categories.

credentials/experience:	'experience', 'certified', 'must have passing score on NTE'
pejorative:	'common sense', 'maturity', 'awareness of what's going on'
too broad/general:	'excellent teacher', 'ability', 'expertise'
miscellaneous:	'leadership', 'family person', 'growth oriented'

The pilot indicated that experience and/or credentials could be anticipated, but we chose to exclude this from the list of traits since it was less relevant to the current effort. In all, 119 respondents mentioned affective traits (C&C, ENT, CIS, and RAP) while 32 mentioned cognitive traits (INT, SAK).

When the number (or, equivalently, the proportion) of respondents choosing a response category was used for scaling the importance of the qualities/categories, the resulting correlations (Table 2) with the Likert, Thurstone and Agree scalings of the same qualities were somewhat disappointing, being only 0.70, 0.51, and 0.72, respectively. The lower correlations appeared to be due primarily to the poor showing of integrity and character, I&C, within the free-response format. In fact, when I&C is removed, the correlations above jump to 0.90, 0.88, and 0.88, respectively.

The explanation would seem to be that when I&C is specifically listed (as with the Likert and Thurstone paired comparisons) I&C is seen to be very important, but when it is not suggested (as with the free-response) I&C tends to be either assumed or overlooked. I&C ranked first using a Likert (or Agree) format and near last using a free-response format. The item format and scaling method were certainly relevant factors in assessing the perceptions of the principals, but by using several approaches the differences could be at least be partially explained.

A Composite Scale

With I&C removed, the scale values of the remaining eight qualities using the Likert, Thurstone, and free-response formats were put on an arbitrary scale (a T-scale, with mean 50, s.d. 10). Linear transformations such as the T are admissible with interval level measures such as the Thurstone and free-response (frequencies are on a ratio scale), but admittedly suspect with only mock-interval scales such as the Likert. The mean T-scale values for each quality over the three methods was selected as a composite scale value. These values are shown in Table 1 under 'Composite' and the correlations with the original scales in Table 2. The 'Agree' scale was not used in computing the composite since it was

a derivation of the Likert scale and not based on a unique set of items. The high correlations (0.95 to 0.97) with all of the original scales are partially inflated due to the correlation of any component with the whole, but give confidence nonetheless in the validity of the measure.

Additional verification of the scale values was done by using item response theory (Hambleton & Swaminathan, 1985), a more modern scaling approach, with the (10) Likert items. The 'graded response' model of Samejima (1969) was fit to the data. For all of the items, estimates of three of the four thresholds for the five response categories (the category for 'Strongly Disagree' had too few responses to be accurately estimated) correlated 0.93 to 0.96 with the composite scale values.

Personal Strengths of the Principal

A series of three Likert items regarding the respondent's (principal's) own perceived affective and cognitive strengths were: 'As a teacher, I feel that my knowledge and understanding of the subject-matter in my field was greater, in general, than my skill in working with children.' (COG>AFF), 'As a teacher, I feel that my skill in working with students was well above average when compared to that of other teachers in my school.' (AFF>AVG), and 'As a teacher, I feel that my subject-area knowledge and understanding were well above average when compared to that of other teachers in my school.' (COG>AVG).

The mean response to COG>AFF was 2.33 (s.d. 0.84) where 2 was 'Disagree' and 3 was 'Neutral'; 71.2% of the respondents disagreed (13.1% agreed) with this statement. The mean response to AFF>AVG was 4.00 (s.d. 0.839) or 'Agree' with 78.9% agreement and 4.4% disagree. There was not one respondent who strongly disagreed with this statement. Similarly, the mean for COG>AVG was 3.49 (s.d. 0.83), between 'Neutral' and 'Agree', with 51.9% agreement, 12.6% disagreement and no respondent selecting 'Strongly Disagree'. The correlation between COG>AVG and AFF>AVG was 0.304 ($p<0.001$). In general, these 271 principals perceived themselves as being somewhat above average in the cognitive dimension and well above average in the affective dimension and about 9% of the variation in one rating could be attributed to variation in the other.

The only significant correlations between any of these three items and the Likert items on the desired characteristics were between COG>AVG and INT ($r=0.267$, $p<0.001$) and COG>AVG and SAK ($r=0.262$, $p<0.001$). That is, a tendency to agree with COG>AVG regarding the principal's personal skills coincided with a slightly increased desirability of both intellectual capacity and subject-area knowledge in a teaching candidate. In fact, the summated Likert score was significantly correlated only with COG>AVG ($r=0.238$, $p<0.001$).

Possible Explanations

There were six Likert items that alluded to possible reasons

for preferring affective to cognitive skills or vice-versa. The items were:

> 'In general, teachers who choose to leave classroom teaching, for whatever reason, tend to have been those with above average scholastic aptitude.' (LEAVE)
>
> 'Teachers for whom academic work comes very easily tend to have more difficulty understanding the learning problems of the less able students.' (PROBS)
>
> 'Teachers with more scholarly interests seem better able to relate to their students.' (RELATE)
>
> 'The most successful teachers in a school are most often not the most intellectually gifted teachers in the school.' (GIFTED)
>
> 'A teaching candidate's undergraduate academic record, if an average of C or better was maintained, should play a minor role in the selection process.' (SHOULD)
>
> 'A teaching candidate's undergraduate academic record, if an average of C or better was maintained, usually does play a minor role in the selection process.' (DOES)

All were used with 'Strongly Agree' scored 1 to 'Strongly Disagree' scored 5 except LEAVE and RELATE which were reverse scored. The responses to each item are shown in Table 3.

Insert Table 3 about here

Overall, there was a tendency to somewhat disagree with LEAVE and RELATE and to somewhat agree with PROBS, GIFTED, and DOES. The responses to SHOULD were bimodal, evenly divided between 'Agree' and 'Disagree'. Significant correlations with the Likert items on the desirable qualities were seen only with SAK (SHOULD, $r=0.237$, $p<0.001$) and with INT (LEAVE, $r=0.222$, $p<0.001$; RELATE, $r=0.290$, $p<0.001$). That is, respondents who valued SAK less tended to agree that GPA should play a minor role; respondents who valued INT less tended to both disagree that the more scholastically able teachers leave the classroom and to disagree with the ability of the more scholarly teacher to relate to the students.

The relationship between SHOULD and DOES, Table 4, was

Insert Table 4 about here

informative. Altogether, there were 172 out of the 266 (64.7%) principals who responded to both items who thought that a teaching candidate's undergraduate academic record (GPA), if 'C' or better, **does** have the role that it **should** have in the selection process. That is, 64.7% of the respondents rated 'should' and 'does' identically. There were 28 (10.5%) who thought that GPA should play a lesser role than it now does and, more than twice as many, 66 (24.8%) who thought that GPA should play a greater role than it now does.

This result is somewhat incongruous with respect to the larger attention paid to the affective attributes of candidates, presumably at the expense of the cognitive attributes. Indeed, when the Likert scaling of the original nine qualities is redone using only the rating of this group of 66 principals for whom GPA was not as important as it should be, SAK (3.74) and INT (3.22) retained their identical (Likert) ranks. It would appear that the affective qualities were most attractive even to the subsample of principals who felt that academic records should play a larger role in the selection process.

The 'Best and Brightest'

To determine if there was a significant preference for or against the 'best and brightest', a brief description of a candidate was rated on a scale from 1 (not at all desirable) to 7 (extremely desirable). Four forms using two levels of GPA (2.6 and 3.6) and two undergraduate universities (Stanford and Western Michigan) were randomly sent to respondents. There were 265 responses to this item from the 271 completed surveys. In addition, this was the single item that the additional 117 original non-respondents had agreed to answer. The analysis was a 2X2X2 analysis of variance. Certainly, the levels of GPA and university were few and arbitrary, but with eight cells in the design and a total of 382 subjects the power (0.84) to detect a 'small to medium' effect (0.15) at alpha = 0.05 was deemed adequate (Borenstein & Cohen, 1988) and would have been reduced to unacceptable levels with more cells in the design. The results are presented in Table 5. Since none of the interactions were

Insert Table 5 about here

significant, the main effects are uncomplicated. Notice that, contrary to our expectations, the difference between the respondent groups was significant and that it is therefore unreasonable to pool the data across these groups. Indeed, this finding limits the

generalizability of our study, illustrates both the need to pursue non-respondents rather relentlessly, and reminds us of the inherent limitations of the mailed-survey. Just why the means (Table 6) for those who chose not to complete the original survey after two requests, but were willing to respond to a single question, should

Insert Table 6 about here

be significantly lower (more critical or less accepting) than the responding group is unclear to the authors. The main effects for GPA and university were both significant. The principals showed a significant preference for the candidate with the higher GPA and for the Stanford graduate. There was no bias against the 'best and brightest' seen with this item; the affectively identical but cognitively stronger candidate was given the higher mean rating.

The Summative Question

The final (Likert-scaled) question on the survey was: 'It is critically important to attract the "best and brightest" persons to careers in teaching.' The mean response was 4.07 (s.d. 0.89), or general agreement. Of the 268 respondents to this question, there was not one who strongly disagreed and only 21 who disagreed at all.

Discussion

One clear message from this study was that 'how' a question is asked may have a large effect on both responses and the possible interpretations of those responses. For the pairwise comparisons, options were presented and the choices were forced. This may be more or less realistic in practice depending on the size of the candidate pool and on the factors taken into consideration by the hiring agent. That is, with a small number of candidates, a principal might well be forced to choose between applicants and the desirable qualities they may possess. However, under all conditions, affective qualities were seen as much more important than cognitive qualities, even amongst the subsample of principals who felt that GPA should play a larger role in the hiring process than it now does.

With a Likert format, options were again presented, but the choices were not forced. This might be most realistic with a larger candidate pool. The analyses showed some distinct differences between the paired comparisons and the Likert items (e.g., neither the first nor last rated qualities were the same); previous research has shown that the difficult decisions required for paired comparisons may result in a different scaling than from an unforced method (Alwin & Krosnick, 1985). It was nevertheless the case that the scales correlated reasonably well (0.83) and that the affective qualities were at the top in both scalings and the

cognitive qualities were at the bottom in both scalings. Neither rescaling the five-option Likert items to a simple agree/neutral-disagree format nor using a more sophisticated approach (item response theory) had a serious effect on the scaling of the qualities.

The most realistic method of determining the relative importance of possible candidate qualities might well have been the free-response format. Unless a principal had a highly structured approach to the selection process, a decision might rest upon his or her perceptions of a candidate's worth with respect to some 'single most important quality'. Since 76.0% of the usable free-responses to this item fell within the list of nine previously identified qualities (86.3% if the purposefully-omitted experience/credentials and the too-general responses are not included), the list of nine traits would seem to be reasonably complete.

The scaling of responses to this item again tended to confirm the general affective preferences of the principals and correlated well (0.88 and 0.90) with the other scaling approaches when 'integrity and character' was omitted. It would seem reasonable to suppose that I&C was 'assumed' rather than 'unimportant' to the majority of principals in the same way that 'lack of a drug problem' (drug use was not mentioned and is only illustrative) was likely assumed, but would certainly not be seen as unimportant.

The results of the various approaches to identifying and measuring the importance of the teaching candidate qualities converged to a composite scale and to the conclusion that cognitive qualities (including the much touted 'subject-area knowledge') are seen as substantially less important than affective qualities at the secondary level. The composite scale would have us believe that intellectual capacity was seen as least important of the eight or nine qualities desired by principals and that relative to intellectual capacity, concern and caring was seen as twice as important as subject-area knowledge. At the elementary level, one might expect even more importance being given to the affective traits, if this is possible.

A methodological caveat is needed in that the responses of the original non-respondents to the entire survey who agreed to answer a single item follow-up question gave significantly different ratings to that item. The pattern of the responses was similar, but the differences limit the generalizability of the current study. In particular, extrapolation to the sentiments of the 111 (22.2%) principals who choose to not participate at all are totally unwarranted.

Are cognitive qualities so unimportant as to be seen as a liability? The experimental item with two levels of undergraduate university and GPA gave a negative answer to this question. There was no bias against the stronger candidate, indeed, there was a significant preference for the more able candidate with respect to both university and GPA. Admittedly, the limitations of this item are legion: the question specifically referred to a social studies

teacher; only levels of GPA (2.6 and 3.6) were investigated; Western Michigan and Stanford are distinct in many unintended ways; single items tend to be substantially less reliable than summated scales. In any event, the question of possible bias now seems somewhat less important in light of the relative lack of concern with cognitive skills.

The principals did see themselves as the hiring agent to a large extent. They also saw themselves as being above average in both the affective and cognitive domains, but with more affective skills than cognitive. Associations between these personal items and the nine qualities were only apparent with the cognitive items where significant positive correlations indicated that when the principal perceived his or her own cognitive skills as stronger, there was a tendency to put a larger value put on the cognitive skills of a teaching candidate. It would be a clear over-conclusion to say that only those principals who perceive themselves as having greater cognitive skills value those cognitive skills in teaching candidates, but there is some indication of a tendency in this direction.

Identifying the reasons for the affective preferences of the principals was largely unsuccessful. Overall, there was slight, but general, agreement (when the appropriate scale reversals are taken into account) with all of our hypothesized old-wive's tales: the best teachers don't leave, bright teachers don't relate well to typical kids, the most successful teachers are most often not the brightest, bright teachers can't understand the problems of less bright students, and GPAs above a 'C' don't, and shouldn't, mean much.

Whether it is the cumulative effect of such nonsense or whether there is some unexamined mechanism at work causing cognitive skills to be devalued remains unanswered. At this point we can only play our findings against our own experiences in the public schools in an attempt to go beyond the research findings in search of an explanation. Our hypothesis focuses on the potency of two different but related needs in the schools. One is a need of students and the other is a need of school personnel.

It is a rather common perception that increasingly large numbers of students today come to school with enormous basic human needs that are often only partially fulfilled. These include needs for nutrition, clothing, shelter, physical and psychological protection, self-esteem, and, in general, concern and caring. That principals and teachers would see meeting these needs to be essential prerequisites to intellectual development is altogether plausible and thus a possible explanation for an emphasis on the affective characteristics of teachers.

To make the argument more compelling, consider a basic need of school personnel, a safe and orderly learning environment. It would seem that schools and their principals most immediately pay the largest public price for any failure to maintain an acceptable level of order and control among students and staff. However, if

the affective needs of the students are not being met, then it seems reasonable that the resulting frustration is apt to be reflected in that very area where schools and their principals are most vulnerable. That is, failure to maintain a safe, orderly, and controlled place to learn would seem most likely to occur when the physical and affective needs of students are being frustrated rather than when the academic or intellectual needs of the students are being frustrated.

While we have reasoned our way to the principals' priority on the affective characteristics of teachers which would help the school organization to meet its own need for self-preservation, we certainly have no illusions that this conjecture is the only possible explanation for our findings. In fact, we are not even certain that the priorities professed by these principals are played-out in practice. But, if they are, the many reform efforts that encourage luring the 'best and brightest' to careers in teaching may be frustrated at the door to the principal's office.

References

Alwin, D. F., & Krosnick, J. A. (1985). The measurement of values in surveys: A comparison of ratings and rankings. Public Opinion Quarterly, 49(4), 535-552.

Baldwin, C. L., Collins, J. R., Kostenbauer, T., & Murphy, C. B. (1988). Research choices for measuring outcome of high school groups. Journal for Specialists in Group Work, 13, 2-8.

Berliner, D. C. (1987). Teacher selection in the Mesa Unified School District. In A. E. Wise et al. (Eds.), Effective Teacher Selection: From Recruitment to Retention, Case Studies (pp 1-51). Santa Monica, CA: The RAND Corporation.

Borenstein, M., & Cohen, J. (1988). Statistical Power Analysis: A Computer Program. Hillsdale, NJ: Lawrence Erlbaum.

Brophy, J. E. (1982). How teachers influence what is taught and learned in classrooms. The Elementary School Journal, 83(1), 1-13.

Dunn-Rankin, P. (1983). Scaling Methods. Hillsdale, NJ: Lawrence Erlbaum.

Gips, C. J., & Bredeson, P. V. (1985). The selection of teachers and principals: A model for faculty participation in personnel selection decisions in public schools. High School Journal. December 1985-January 1986.

Gips, C. J., & Johanson, G. A. (1989, October). Principal's priorities for teacher characteristics in employment decisions. Paper presented at the meeting of the Midwest Educational Research Association, Chicago, Il.

Hambleton, R. K., & Swaminathan, H. (1985). Item Response Theory: Principles and Applications. Boston, MA: Kluwer-Nijhoff.

Henry, M. A., & Sa'ad, F. F. (1977). The impact of field experience of student teachers' perceptions of good and poor teaching. The Teacher Educator, 13(2), 18-22.

Johnson, S. T., & Prom-Jackson, S. (1986). The memorable teacher: Implications for teacher selection. Journal of Negro Education, 55(3), 272-283.

Kowalski, T. J., & Weaver, R. A. (1989, February). Characteristics of outstanding teachers: An academic and social involvement profile. Paper presented at the meeting of the Association of Teacher Educators, St. Louis, MO.

Leinhardt, G., & Greeno, J. G. (1986). The cognitive skill of teaching. Journal of Educational Psychology, 78(2), 75-95.

Provalko, R. M. (1970). Recruitment to teaching: Patterns of selection and retention. Society of Education, 43(3), 340-355.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph Supplement, #17.

Sandefur, J. T., & Adams, R. A. (1976). An evaluation of teaching: An interim research report. Journal of Teacher Education, 27(1), 71-77.

Schlecty, P. C., & Vance, V. S. (1981). Do academically able teachers leave education? Phi Delta Kappan, 62(10), 106-112.

Weaver, W. T. (1979). In search of quality: The need for talent in teaching. Phi Delta Kappan, 46(9), 29-32.

Table 1
Scale Values of the Nine Teacher Characteristics

	Scaling Method				
Quality	Thurstone	Likert	Free-Response	Agree	Composite
C&C	1.27	4.36 (0.61)[a]	0.18 (36)[b]	0.97	61.51
ENT	0.96	4.25 (0.66)	0.15 (29)	0.93	55.85
I&C	1.56	4.24 (0.76)	0.06 (12)	0.91	-
CIS	1.38	4.22 (0.59)	0.13 (26)	0.95	57.56
C&F	0.92	4.12 (0.70)	0.12 (23)	0.91	52.33
RAP	0.87	4.10 (0.74)	0.14 (28)	0.90	53.59
SAK	0.75	3.64 (0.78)	0.12 (23)	0.65	47.12
DED	0.00	3.58 (0.79)	0.05 (10)	0.61	36.53
INT	0.31	3.23 (0.86)	0.05 (09)	0.41	35.51

[a]standard deviation
[b]number of respondents

Table 2
Correlations Between the Scale Values of the Nine (Eight) Teacher Characteristics

	Scaling Method				
Method	Thurstone	Likert	Free-Response	Agree	Composite
Thurstone	-	(0.84)	(0.88)	(0.84)	(0.95)
Likert	0.83	-	(0.90)	(0.99)	(0.96)
Free Response	0.51	0.70	-	(0.88)	(0.97)
Agree	0.81	0.99	0.72	-	-

Note. Numbers in parentheses are correlations with I&C deleted.

Table 3
Frequencies (Percentages) for the Items Relating to Possible Reasons for Affective Preferences in Hiring Decisions

	Response				
Item	S Agree	Agree	Neutral	Disagree	S Disagree
LEAVE	2(00.7)	79(29.5)	85(31.7)	92(34.3)	10(03.7)
PROBS	16(05.9)	134(49.8)	39(14.5)	71(26.4)	9(03.3)
RELATE	1(00.4)	30(11.2)	90(33.5)	141(52.4)	7(02.6)
GIFTED	39(14.5)	139(51.7)	48(17.8)	39(14.5)	4(01.5)
SHOULD	6(02.2)	109(40.5)	47(17.5)	102(37.9)	5(01.9)
DOES	6(02.3)	138(51.9)	48(18.0)	72(27.1)	2(00.8)

Note. Percentages may not sum to 100 due to rounding.

Table 4

Crosstabulation of SHOULD (GPA should play a minor role) By DOES (GPA does play a minor role)

	DOES					
SHOULD	SA	A	N	D	SD	Row Total
SA	3	2	1			6
A	2	85	8	13		108
N		14	29	4		47
D	1	36	10	53		100
SD		1		2	2	5
Column Total	6	138	48	72	2	266

Table 5
Factorial Analysis of Variance of Candidate Desirability with GPA (2) X University (2) X Respondent Group (2)

Source of Variation	Sum of Squares	DF	Mean Squares	F	p
Main Effects	93.615	3	31.205	23.187	.000
University	12.838	1	12.838	9.539	.002
GPA	69.407	1	69.407	51.574	.000
Respondent Group	6.925	1	6.925	5.145	.024
Interactions					
Univ X GPA	.084	1	.084	.062	.803
Univ X Group	.196	1	.196	.146	.703
GPA X Group	.428	1	.428	.318	.573
Univ X GPA X Group	.031	1	.031	.023	.879
Explained Variation	94.375	7	13.482	10.018	.000
Residual Variation	503.324	374	1.346		
Total	597.699	381	1.569		

Table 6
Cell Means, Marginals, and (Count) for the Single Candidate Desirability Item

	Respondent Group		
Group	Entire Survey	Single-Item	
Stanford			Total
High GPA	5.64 (72)	5.20 (25)	5.53 (97)
Low GPA	4.70 (60)	4.45 (29)	4.62 (89)
Stanford Total	5.21 (132)	4.80 (54)	5.09 (186)
Western Michigan			
High GPA	5.20 (64)	4.90 (31)	5.11 (95)
Low GPA	4.35 (69)	4.16 (32)	4.29 (101)
Western Michigan Total	4.76 (133)	4.52 (63)	4.68 (196)
Total High GPA	5.43 (136)	5.04 (56)	5.32 (192)
Total Low GPA	4.51 (129)	4.30 (61)	4.44 (190)
Total	4.98 (265)	4.65 (117)	4.88 (382)