Effects of a State Mandated Policy (Site-Based Councils) and of Potential Role Incumbents on Teacher Screening Decisions in High and Low Performing Schools

I. Phillip Young
University of California-Davis

Kimberly Miller-Smith
Southwest Licking County Local School District-Ohio


Abstract
Some states have viewed teacher selection as a means of improving student performance and have mandated the use of site-based teacher councils. To assess the utility of this legislative action, an experimental study was conducted. This study uses a 2X3X2 factorial design that varies state legislation, role of the decision maker, and academic performance of the school site. Credentials of hypothetical teacher candidates were evaluated as if screening for a vacant position, and evaluations were submitted to a MANOVA. Results indicate that legislated alterations in the teacher-selection process failed to have any substantial effects on outcomes in the screening of teacher candidates for elementary school positions. Although screening decisions were found to be the same for participants affiliated with both low- and high-performing school districts, teachers were more appreciative of candidate’s credentials than either principals or parents.

Keywords: teacher selection; screening decisions; site-based councils.
Introduction

Importance of teacher selection is well documented within the professional literature. As early as the 1930’s Graces (1932) noted that “wise selection is the best means of improving the system, and the greatest lack of economy exists whenever teachers have been poorly chosen or inadequately adapted to their profession” (p. 191). This sentiment, as noted by Graces, continues to be echoed today (Chounet, 2002). “In recent years, the issue of teacher quality has once again risen to the top of the school reform agenda” (Liu & Johnson, 2003).

Given the importance of teacher selection, the process by which teachers are selected to fulfill their assigned duties within the public school setting varies across states of the Union (Liu & Johnson, 2003). Most states have and continue to delegate this responsibility to educational administrators (Cuban, 1988), while at least one state (Kentucky) vests this authority with site-based councils (SBC) (see KERA, 1990). Potential differences in these legislative actions relative to the screening and selection of teachers have failed to receive any consideration within the professional literature. As a result of this void it is unclear, to date, if legislative actions mandating SBCs have any effects on outcomes for the teacher-selection process.

Such information has important implications for guiding public policy makers when shaping the teacher-selection process at the state and local school district level. To provide empirical data about teacher-selection processes differing according to legislative intent is the purpose of this study. Within this study, we explore this thorny issue of different configurations for the selection of teachers from several different perspectives.

First, we examine a particular type of selection decision. The type of selection decision considered in this study involves screening applicants on the bases of their paper credentials. Well noted in the selection literature is that “there are situations where decisions are made on the basis of paper information” (Cleveland & Landy, 1983, p. 619) submitted by applicants. According to research from the private sector (Macon & Dipboye, 1994; Thoms, McMasters, Roberts, & Dombkowski, 1999), as well as the public sector (Liu & Johnson, 2003), decisions made on the basis of paper credentials serve a gate keeping function for delimiting an initial applicant pool. As indicated by Cable and Gilovich (1998), “because interviews are time intensive and costly, interviewers generally prescreen job applicants on the basis of their resumes before granting interviews” (p. 501).

Second, we hold constant the type of selection decision explored in this study (screening as opposed to interviewing decision) and vary the legislative actions of state governments. Legislative actions of state governments are manipulated to include one state (Kentucky, experimental condition) requiring SBCs and to include another contiguous state (Ohio) failing to delegate through legislative processes SBCs as a quasi control (see Campbell & Stanley, 1963). By comparing (main effect) screening decisions for teacher candidates between these states (experiment vs. control), we provide certain insights about the potency of legislative actions on the part of state legislatures relative to a particular phase of the teacher-selection process.

Third, potentially confounding any differences for a main effect involving legislative intent on the part of state governments varying in methods for selecting teachers between states are the type of potential participants taking part in the selection process. That is, if a main effect were detected for the legislative process, then is this main effect due either to differences from the legislative enactments of state governments or to varying perceptions of different roles of participants? To sort among these possible explanations, we manipulated also the role of potential participants between both states in this study.
Fourth, we explore the outcomes of teacher screening decisions for high-performing and low-performing school districts. Implied by existing teacher selection research is that high-performing school districts may do something different from low-performing school districts within the teacher-selection process (Wise, et.al., 1987). To assess if public school districts in general should follow those practices used by high-performing school districts from a modeling perspective is still another purpose of this study. More pointedly, the research questions addressed in this study are as follow: (1) Are teacher screening decisions different in a state mandating SBCs for selection from screening decisions in a state failing to mandate SBCs; (2) are teacher screening decisions a function of the roles held by potential SBC members; (3) are teacher screening decisions related to the performance of a public school district; and (4) are teacher screening decisions related systematically to some combination of state legislation, role of potential incumbents, and/or performance of a school district?

Literature

According to the emerging body of teacher selection research, the teacher-selection process includes decisions of individuals as applicants and decisions of administrators as employers (Winter, Ronau, & Munoz, 2004). Both research streams contribute to knowledge about the recruitment and selection of teachers but in varying ways. Research addressing these different decision sources (applicants and administrators) has produced two separate but related research streams. One research stream emerges from the recruitment literature and focuses on the decisions of applicants (for a recent review see Winter & Melloy, 2004), while the other stream associated with the selection literature addresses the decisions of organizational representatives (discussed below). It is the later research stream that is the focus of our study, the decisions of organizational representatives as potential employers within the teacher-selection process.

Decisions of organizational representatives are purported to occur at separate stages within the teacher-selection process. One stage involves screening decisions where the initial applicant pool is reduced to include only those candidates perceived to be worthy of further consideration within the teacher-selection process (Young & Ryerson, 1986). The other stage involves interviewing decisions to determine which candidates will receive actual job offers at the culmination of the selection process (Young & Delli, 2002).

Research indicates that screening and interviewing decisions differ in another significant way beyond outcome of the selection process (Gorman, Clover, & Doherty, 1978), and this difference demands separate consideration within the research literature for understanding the teacher-selection process. Importantly, these decisions differ with respect to the mode of applicant stimuli on which each type of decision is based (Macan & Dipboye, 1994).

That is, screening decisions are based on paper credentials submitted by teacher applicants (for a model, see Young & Delli, 2002). On the other hand, interviewing decisions are based on the interpersonal interactions between organizational representatives and applicants (Cable & Gilovich, 1998). Because of these important differences in outcomes (screening versus hiring) as well as in applicant stimuli (paper credentials versus interpersonal interactions) driving selection decisions, we focus in this manuscript on screening as opposed to interviewing decisions.

To investigate screening decisions of organizational representatives within the teacher-selection process, empirical studies have been characterized largely as being driven by a legislative enactment and/or interpreted through the lens of a particular person perception theory. Most legislative studies are couched within the context of specific federal enactments, while most theory-based studies are rooted in a particular person perception theory. Each of these orientations
(legislative and theory) contributes in unique ways to the general knowledge about the teacher-selection process and warrants some attention in this manuscript.

Legislative Studies

Specific federal enactments receiving attention within the teacher selection literature include the Age Discrimination in Employment Act (ADEA), Title VII of the Civil Rights Act, and/or the American with Disabilities Act (ADA). Within legislative studies, characteristics of hypothetical teacher candidates are manipulated to yield applicants either covered by or exempted from the legislative acts as noted. That is, hypothetical teacher candidates are depicted under or over 40 years of age as specified by the ADEA, being female or male as addressed by Title VII of the Civil Rights Act, and/or having acknowledged or unacknowledged handicapping conditions as covered by ADA. Reactions of principals toward these different configurations of applicants have been assessed relative to compliance with federal guidelines. With respect to chronological age of teacher candidates as addressed by the ADEA, screening decisions for hypothetical teacher applicants have been assessed in several studies. Young and Allison (1982) varied the chronological age of teacher candidates and found that both superintendents and principals preferred 29 year old teacher candidates over 49 year old teacher candidates at the screening stage of the teacher-selection process. Several other studies have reported similar evidence of age discrimination at the screening stage of the selection process (Young & Joseph, 1989; Young & McMurray, 1986; Young & Voss, 1986).

As with chronological age of teacher candidates as covered by ADEA, sex of teacher candidates has been explored at the screening stage of the teacher-selection process relative to compliance with Title VII of the Civil Rights Act. Stollard (1990) as well as Wallich (1984) failed to detect any preference for a particular sex group. However, Reis, Young and Jury (1999) found that participants preferred females over males at the screening stage of the selection process.

Only a single study found in the published literature addresses the compliance of public school administrators with the Americans with Disabilities Act. Young and Prince (1999) varied acknowledged disability of teacher candidates under three experimental conditions: confined to a wheel chair, limited to a crutch, or no disability [the control condition]. Their results indicated that disabled teacher candidates were more likely to be extended an interview offer than teacher candidates without an acknowledged disability.

Person Perception Theories

In contrast to legislative studies addressing the compliance of public school administrators, most theory-based studies focus on general psychological/sociological theories pertaining to person perceptions. General person perception theories drawn from the psychological/sociological literature have been used to cast the reactions of organizational representatives about teacher candidates on a theoretical framework. Goals of theory-based studies within the teacher-selection context are to provide a framework for understanding the dynamics of the decision making process from a broad perspective (Arvey & Faley, 1992) beyond statistical effects and are not to provide specific tests for particular theoretical orientations (Young, Chounet, Buster, & Sailor, 2005). That is, theory-based studies guide the interpretation of outcomes rather than provide confirmations of theoretical constructs for testing theories.
Theory-based studies within the teacher-selection context have approached this administrative task involving the screening of teacher candidates from several perspectives as a framework for understanding the decision making of public school administrators. Decisions made by public school administrators have been cast onto several different theoretical frameworks. Included among these frameworks are decremented theory, similarity-attraction theory, and social distance theory.

The decremented theory (Botwinick, 1978) pertains to the aging process and suggests that younger job candidates will be more suited for high activity positions than older job candidates. To examine the utility of this theory for explaining teacher selection decisions, Young, Rinehart, and Baits (1997) varied the physical-demand characteristics associated with particular focal teacher positions in two separate analyses. Physical-demand characteristics varied by Young, Rinehart, and Baits include focal positions between disciplines (physical education vs. physics) as well as within a discipline (general science vs. chemistry). They found that younger teacher applicants were preferred over older teacher applicants for high physical demand focal positions.

Similarity-attraction theory (Byrne, 1961) indicates that organizational representatives (public school administrators) would recommend more likely those similar to self for an interview than those dissimilar to self. Similarity has been explored relative to age cohorts (Bowman, 1999) and to sexual groupings (Reis, et al., 1999) within the teacher screening process. Little support has been reported for this particular theoretical orientation as a framework for understanding teacher screening decisions within the public school setting.

In contrast to the similarity-attraction theory, social distance theory (Bogardus, 1967) attempts to explain how individuals (organizational representatives) view those different from themselves. Research in this area has held constant characteristics of the organizational representative and has varied characteristics of applicants. Several studies (Young & Fox, 2002; Young & De La Torre, in press; Young & Oto, 2004) have held constant characteristics of the organizational representative and have varied national origin of the applicant: Asian candidate, Hispanic candidate, or Native American candidate. Results from these studies indicate that national origin of job candidates moderate screening decisions of job candidates for different focal positions.

**Purpose of this Study**

Results of teacher selection research following these methodological approaches involving either legislative studies or theoretical frameworks have considerably advanced current knowledge about this important administrative task within the educational setting. As noted in our review of published teacher selection studies, most legislative studies explored screening decisions from an outcome perspective (do administrators conform?) relative to various protected class statuses of applicants as defined by existing federal enactments (age, disability status, sex, and ethnic group). Absent from this literature is any study of mandated initiatives bearing on the structure of the selection process per se. One such state initiative that has altered substantially the structure of the teacher-selection process is the Kentucky Education Reform Act (KERA) (1990), and we examine this alteration for teacher selection from several perspectives.

KERA decentralizes the teacher-selection process from a unilateral procedure involving solely administrators to a multilateral procedure involving multiple stakeholders. This change is accomplished through the use of site-based teacher councils comprised of principals, teachers, and parents (Lindle, 2000). If this particular state’s legislated initiative involving site-based decision councils has any effect on the outcomes of the teacher process, then we would expect screening decisions for teacher candidates to be different from those assessed in a state not mandating site-
based decision making. To assess potential effects associated with different teacher-selection processes, we compare screening decisions made in Kentucky mandating site-based decision councils with screening decisions made in a contiguous state (Ohio) vesting teacher screening decisions legislatively with school administrators.

Within our study, we commingle legislative actions with a theory-based framework. Theory-based studies addressing teacher selection have approached this administrative task from an interactional perspective through varying personal characteristics of administrators and applicants to create similar and dissimilar pairings. In contrast to existing research protocol, we hold the characteristics of applicants constant and varied the role of potential evaluators (principal, parent, or teacher). The role of the evaluator may well influence outcomes of the selection process. According to role theory, an individual’s role within the teacher-selection process is an important determinant for selection outcomes because role shapes perceptions of decision makers beyond context (Michener, DeLamater, & Myer, 2004). If role theory has any utility within the present context as a theoretical framework for explaining screening decisions made by organizational representatives, then we would expect to detect either a main effect for role of the evaluator or an interaction effect between role of the evaluator and legislative actions of states varying the teacher-selection process (site-based vs. non-site-based).

Results from state accountability measures indicate great variability among public schools relative to performance on academic performance index measures. It is well noted that some of this variability in performance is due, at least in part, to the proficiency of the classroom teacher (Chounet, 2002; Graces, 1932), and it is assumed generally that high-performing schools do something different from low-performing schools within the teacher-selection process. This assumption has deep roots with a seminal study advocating the selection practices of only purported flagship school districts (Wise, et.al., 1987) as a means of improving the selection process and advancing the learning of students. Although this study by Wise, et. al. (1987) is dated, it is important to note that the general assumption of high-performing school districts doing something different from low-performing school districts within the teacher-selection process has escaped entirely empirical investigation within this important area. To assess for potential differences between high-performing and low-performing school districts relative to the teacher screening process is another advancement of this study. If this basic modeling assumption has any validity, then we would expect that high-performing schools would make teacher screening decisions different from low-performing schools.

Method

The population for this study is all public elementary school buildings in two neighboring states, Kentucky and Ohio. Kentucky mandates site-based teacher councils for the selection of teachers, while Ohio vests this authority with public school administrators. Given this particular population configuration, consideration was afforded to the power of statistical tests and to potential return rates for selecting participants to take part in this study.

A power analysis was performed for a fixed effect 2X3X2 factorial design with a conventional medium effect size (Ω² = .25), a tradition statistical power of .80, and an established alpha level of .05 (Cohen, 1977). This calculation rendered a total sample size of 168 participants with 14 participants per cell. A review of existing return rates from similar types of studies suggested over-sampling this population to insure adequate statistical power (Newton, Giesen, Freeman, Bishop, & Zeitoun, 2003), and a total of 624 participants were sampled at random and requested to participate in this study.
Independent Variables

One independent variable manipulated in this study is the state-legislated process for selecting teachers. The state-legislated process was varied to include elementary schools mandated to use site-based councils and elementary schools exempted from such legislation. To operationalize these two different experimental conditions, half of our sample was selected at random from a state delegating teacher selection to site-based councils (Kentucky) and half of our sample was selected at random from a state delegating teacher selection to administrators (Ohio).

Another independent variable manipulated in this study is the role of the evaluator within the teacher screening process. The role of the evaluator was varied to include principals, teachers, and parents as potential legislated site-based council members identified by KERA. To increase the statistical integrity of our analysis, several steps were taken in the selection of these different role incumbents across both states.

When selecting potential role participants for taking part in this study, we were concerned about independence of observations, basic knowledge about school operations, and motivational level for participants. With respect to independence of observations, characteristics of elementary schools vary in many ways (size, diversity, faculty demography, instructional programs, school climate, etc.), and each of these ways or some combination of these ways could influence screening decisions at the local school building level if more than a single role participant is selected from the same school. To control for a dependency within as well as across buildings and states, only a single role incumbent (principal, teacher, or parent) was selected at random from any particular school building and requested to take part in this study.

In addition to being concerned about dependency among participants within the teacher-selection process, SBC members in Kentucky are involved with and are exposed to basic information about several other areas in the performance of their duties. Examples of some of these areas are space utilization, discipline, classroom management, and extra curricular programs (see Newton, Winter, & Keedy, 2001). Although this type of information is available readily to principals in general and teachers in general when fulfilling their assigned job responsibilities, such is not the case always for parents, especially a random sample of parents in general likely containing many of those detached by choice from school operations.

Also, a random sample of parents in general would likely yield a sample of potential participants for this study unwilling to assume any leadership responsibilities as required for SBC membership. To control for knowledge about school building operation as well as motivation to serve in a school building leadership capacity within as well as between states, we selected at random from PTA/PTO presidents. This choice, as a means for operationalization of the parent role in our study is far from perfect but perhaps the best among alternatives in this type of legislative study. This choice of PTA/PTO presidents controls at least partially for both basic knowledge about school operation and willingness to serve in a leadership role at the school building level. All PTA/PTO presidents are parents even though all SBC council members are not PTA/PTO presidents. Most importantly, an exact match between parents fulfilling SBC membership roles in Kentucky fails to exist in our control state (or any other state), but PTO/PTA presidents exist in both states, and these parents are likely to share a similar knowledge base about school operations and to have assumed a leadership role by fiat at the local school building level.

Still another independent variable manipulated in this study was the academic performance of public school districts, in this study categorized as low vs. high through the median split on a composite accountability measure. Contained within this accountability measure was student performance on state proficiency tests, promotion rates, and attendance data. Because each state
used a different proficiency test for assessing academic performance, all competency measures (proficiency test, promotion rates, and attendance data) were transformed to a standardized score (z score) within each state. Standardized scores were then summed to compute an overall composite score (accountability indices) within each state. Within each state, a median split was performed on the basis of accountability indices based on composite measures encapsulating academic performance, graduation rates, and attendance standards.

This process has certain disadvantages. The most notable disadvantage is that a much stronger manipulation of low and high performance could have been obtained through using only elementary schools plus or minus one standard deviation above or below the mean accountability measure. This stronger manipulation based on plus or minus one standard deviation involves excluding approximately 68% of the elementary schools in each state (the majority) but strengthens considerably the potency of the manipulation. As such, the research question changes from comparing low- and high-performing school districts to comparing the highest of the high with the lowest of the low.

However, our choice of using a median split has certain advantages. Most notably, by including all elementary schools, our findings generalize to a much broader population. This broader population is likely to have far greater appeal to legislators as well as other policy makers that are interested in what are the best practices for the most and not a select few.

Also, not to be overlooked with our process for defining low versus high-performing elementary schools is the manner by which we performed the median split. We performed our median split on a linear composite score distribution containing multiple measures (academic, promotion and attendance) rather than on a univariate score distribution involving only a single variable for defining academic accountability. Advantages of a linear composite score distribution over a univariate score distribution were noted early on by Fisher (1936), and graphically, this same general (albeit not specific) phenomenon is illustrated more recently by Stevens (2002) using a discriminant analysis example (see p. 303, Figure 7.3).

### Dependent Variables

The two dependent variables focused on evaluations of hypothetical teacher candidates by participants. One dependent variable measured a candidate’s perceived skill levels according to specific job related criteria. The other dependent variable assessed the willingness of evaluators to consider the teacher candidate for a selection interview.

Perceived skill levels of teacher candidates were measured according to six separate criteria. These criteria were the knowledge of curriculum, the ability to transmit knowledge, contribution to the overall school program, the ability to create a friendly classroom environment, the ability to maintain a discipline classroom, and potential for professional growth. Each criterion was rated by participants on a 4-point Likert type scale where a higher rating corresponded to a more favorable evaluation on any specific criterion than a lower rating on the same criterion. Ratings on these criteria were used to calculate an overall composite score for each hypothetical teacher candidate. A coefficient alpha for such a composite score has been reported by others to be .91 (Young & Fox, 2002). Using a 3-week interval for a test-retest procedure, Wallich (1984) reported a stability coefficient of .88 for ratings on these same criteria.

In addition to evaluating the perceived competency of teacher candidates according to specific criteria, evaluators were requested to indicate their probability of extending an interview opportunity to the hypothetical teacher candidate. Probability of an interview offer was rated on a 10 point Likert type scale where higher ratings indicated a greater probability of an interview offer than
lower ratings on this 10 point scale. Lower bound reliabilities for this scale, as assessed with procedures advocated by Wanous and Reichers (1996), are reported consistently to range within the mid .80's within the published literature (Young & Chounet, 2003; Young & Fox, 2002).

Procedure

To assess the effects of independent variables on dependent variables in this study, hypothetical teacher candidates were created through the use of paper type credentials as found in most college placement files and as used by most public school districts (Lui & Johnson, 2003; Thoms, et. al., 1999). Because market demand characteristics can vary greatly according to the type of teacher position sought and influence screening decisions, we held constant the type of focal position sought in this study (Young, et. al. 1997). The particular focal teacher position used in our study is an elementary school teacher.

Included within the paper credentials for each potential elementary school teacher was salient information as found on most typical resumes used by teacher candidates when seeking employment with a public school district. This information included specific sections of the resume addressing in detail career objectives, educational obtainment, professional experiences, extracurricular activities, certification information, professional memberships, community activities, and future ambitions of teacher candidates. To control for potential confounding effects within the credentials associated with age of candidates (Young & Fox, 2002) and sex of candidates (Reis, Young, & Jury, 1999), all dates were omitted and only initials were used for given names.

For each of the experimental conditions, all candidate information was held constant, and each potential participant received an identical set of credentials for hypothetical teacher candidates in the mail. Accompanying the credentials of teacher candidates were a cover letter, a questionnaire for participants, a teacher candidate evaluation form, a pre-addressed envelope for return of information, and a pre-addressed post card for receiving results of the study. Contents of the cover letter solicited the participation of subjects, requested that the individuals evaluate the teacher candidate as if screening applicants for a vacant elementary school teacher position, assured evaluators of anonymity relative to their participation in this study, and offered a complete debriefing if they returned the pre-addressed post card.

After our initial mailing, all participants received a post card as a reminder for their participation. Following the post cards and a time lapse, e-mails were sent to encourage participation in this study. Finally after the post cards and the e-mails, all participants received information packets again by mail and were encouraged to respond if they had not provided already the requested information. Of the 306 participants returning packets, 41 of the participants either failed to follow directions or submitted incomplete information relative to candidate evaluations, and responses from this group were deleted from our data base, leaving a response rate of 49%.

---

1. The reliability of a single item scale as suggested by Wanous and Reichers (1996) represents purportedly a lower bound reliability that builds on the assumption that the correlation \( r_{xy} \) is restricted by the reliability of \( r_{xx} \) (composite evaluation measure) and the reliability of \( r_{yy} \) (probability of an interview offer). Given that \( r_{xx} \) is provided by coefficient alpha (composite evaluation measure) and that \( r_{xy} \) is assessed with the data (correlation between composite evaluation and an interview offer), \( r_{yy} \) (reliability for an interview offer), a lower bound for reliability is calculated through solving for the unknown value relative to reliability being a necessary but not sufficient condition.\\( \hat{r}_{xy} = \frac{r_{xy}}{\sqrt{r_{xx} \cdot r_{yy}}} \).
Design and Analysis

This study used a completely crossed 2X3X2 factorial design where state legislative actions (site-based vs. non-site-based), role of potential evaluators (principal, teacher, or parent), and academic accountability of the elementary school buildings (low vs. high) varied. Because other studies have reported response rates to vary by treatment conditions in similar teacher selection studies and to produce a treatment by response bias (Young & Chounet, 2003), we explored this possibility with our data. A chi-square test was calculated to determine if a systematic relationship existed between response rates and treatment conditions, and results of the chi-square test ($X^2 = 17.88$, df=11, $p<.05$) indicated that a systematic relationship failed to exist between response rates and treatment conditions in our study.

Even though a systematic relationship failed to emerge between response rates and treatment conditions with these data, unbalanced designs, according to Stevens (2002) “is a fairly common occurrence in certain areas of research, and there is no simple solution for this problem” (p. 33). Several solutions are proposed, however, within the methodological literature. Included among these solutions are using an unweighted means analysis (Winer, 1962), a mean substitution procedure (Kirk, 1968), and a sub sampling process (Zar, 1984).

Given that the power analysis, as suggested by Cohen (1977), indicated a need for only 168 subjects to meet our initial parameter specifications, we sampled at random from those completed responses provided by participants. To obtain a completely balanced experimental design, 14 participants were selected at random for each experimental condition to net a final sample size of 168. This latter sampling process for unbalanced designs is recommended by Zar (1984) to obtain a completely balanced design for purposes of statistical analyses.

Prior to submitting our data to a multivariate analysis of variance, several different assessments were performed. Reliability assessments were performed for each dependent variable, with an alpha of .89 for the composite score of ratings summed across the six criteria. With respect to the single item pertaining to the probability of an interview offer rated on a 1–10 point scale, procedures advocated by Wanous and Reichers (1996) suggested a lower bound reliability of .82 with these data.

Given that the MANOVA procedure assumes a multivariate normal distribution and equal variance-covariance matrices, we assessed how well our data met these assumptions. Although a direct multivariate test for normality fails to exist, we assessed for departures in kurtosis and skewness for each dependent variable and found that both the composite scores based on criteria ratings (kurtosis=.79 and skewness=-.71) and the probability of interview offers (kurtosis=.69 and skewness=-.87) met normality assumptions. Likewise, Levene’s test for homogeneity of variance-covariance across dependent variables indicated non-significant differences for the composite scores based on criteria ratings ($F=1.45$, $p=.16$) and the probability of interview offers ($F=.90$, $p=.55$).

Means and standard deviations for the composite ratings of candidates’ perceived proficiency on the job related criteria and the probability of an interview offer for these candidates are found in Table 1, and these data were submitted to a multivariate analysis of variance (MANOVA). Specific tests were performed for each main effect and for all interaction effects. To evaluate these MANOVA effects, a Pillai’s Trace was used, and a statistically significant multivariate main effect was detected only for the organizational role of the participant (Pillai’s Trace=.066, $F(2,156)$, $p=.033$, $\eta^2=.033$) (see Table 2).
Table 1

Summary Statistics for Evaluation Scores by State, Organizational Role, and School Performance

<table>
<thead>
<tr>
<th>School performance</th>
<th>Site Based Legislation (KY)</th>
<th>No Site Based Legislation (OH)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parent</td>
<td>Teacher</td>
</tr>
<tr>
<td>Composite criteria ratings (1–4 scale)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (S.D.)</td>
<td>2.83 (0.39)</td>
<td>2.99 (0.41)</td>
</tr>
<tr>
<td>N</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Low Performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (S.D.)</td>
<td>2.93 (0.59)</td>
<td>3.06 (0.70)</td>
</tr>
<tr>
<td>N</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Likely interview offer (1–10 scale)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (S.D.)</td>
<td>6.86 (1.41)</td>
<td>7.79 (0.37)</td>
</tr>
<tr>
<td>N</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Low Performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (S.D.)</td>
<td>8.00 (1.80)</td>
<td>7.43 (0.24)</td>
</tr>
<tr>
<td>N</td>
<td>14</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 2

Multivariate Analysis of Variances Results

<table>
<thead>
<tr>
<th>Classification</th>
<th>Pillai’s Trace</th>
<th>F Value</th>
<th>F Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reform State Status (A)</td>
<td>.026</td>
<td>2.048</td>
<td>.132</td>
</tr>
<tr>
<td>Organizational Role (B)</td>
<td>.066</td>
<td>2.653</td>
<td>.033*</td>
</tr>
<tr>
<td>School Performance (C)</td>
<td>.031</td>
<td>2.476</td>
<td>.087</td>
</tr>
<tr>
<td>State X Role</td>
<td>.012</td>
<td>0.458</td>
<td>.767</td>
</tr>
<tr>
<td>State X School Performance</td>
<td>.003</td>
<td>0.232</td>
<td>.793</td>
</tr>
<tr>
<td>Role X School Performance</td>
<td>.016</td>
<td>0.643</td>
<td>.632</td>
</tr>
<tr>
<td>AxBxC</td>
<td>.031</td>
<td>1.222</td>
<td>.301</td>
</tr>
</tbody>
</table>

*p < .05, effect = \( \eta^2 = .033 \)

To explore this significant multivariate main effect further as suggested by Weinfurt (2001)—“by far the most popular way of proceeding from a significant effect in MANOVA is to perform univariate ANOVAs for each dependent variable” (p. 262)—we examined separate univariate analysis of variance (ANOVA) for the role of participants relative to each dependent variable. These analyses are found in Tables 3–4. Here, the significant multivariate effect is likely attributable to differences detected for the organizational role of participants relative to ratings on the composite evaluation score, and no statistically significant differences were detected for the probability of an interview offer.
Table 3
Univariate Test Results for Composite Score

<table>
<thead>
<tr>
<th>Classification</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reform State Status (A)</td>
<td>1</td>
<td>0.041</td>
<td>0.000</td>
</tr>
<tr>
<td>Organizational Role (B)</td>
<td>2</td>
<td>1.470</td>
<td>4.841*</td>
</tr>
<tr>
<td>School Performance (C)</td>
<td>1</td>
<td>0.799</td>
<td>2.629</td>
</tr>
<tr>
<td>State X Role</td>
<td>2</td>
<td>0.246</td>
<td>0.810</td>
</tr>
<tr>
<td>State X School Performance</td>
<td>1</td>
<td>0.051</td>
<td>0.167</td>
</tr>
<tr>
<td>Role X School Performance</td>
<td>2</td>
<td>0.153</td>
<td>0.505</td>
</tr>
<tr>
<td>AxBxC</td>
<td>2</td>
<td>0.236</td>
<td>0.776</td>
</tr>
</tbody>
</table>

*p < .05, effect= $\eta^2 = .06$

Table 4
Univariate Test Results for Probability of Interview Offer

<table>
<thead>
<tr>
<th>Classification</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reform State Status (A)</td>
<td>1</td>
<td>6.482</td>
<td>1.664</td>
</tr>
<tr>
<td>Organizational Role (B)</td>
<td>2</td>
<td>6.256</td>
<td>1.606</td>
</tr>
<tr>
<td>School Performance (C)</td>
<td>1</td>
<td>19.339</td>
<td>4.964*</td>
</tr>
<tr>
<td>State X Role</td>
<td>2</td>
<td>2.911</td>
<td>0.747</td>
</tr>
<tr>
<td>State X School Performance</td>
<td>1</td>
<td>1.720</td>
<td>0.442</td>
</tr>
<tr>
<td>Role X School Performance</td>
<td>2</td>
<td>0.339</td>
<td>0.087</td>
</tr>
<tr>
<td>AxBxC</td>
<td>2</td>
<td>8.756</td>
<td>2.247</td>
</tr>
</tbody>
</table>

*p < .05. Although a statistically significant univariate effect is detected for district performance, a multivariate statistical effect was not detected and the later should not be interpreted.

Insight about these differences as detected with the composite criteria ratings is provided by conducting pairwise comparisons among the different organizational representatives (see Table 5). Pairwise comparisons among the different role incumbents revealed two statistically significant differences. Both elementary school principals and elementary school parents differed from elementary school teachers in their perceptions of teacher candidates as assessed on the composite score measuring competencies according to the six criteria used to evaluate teacher candidates.

Table 5
Bonferroni Test of Multiple Comparisons among Role Members on Composite Scores

<table>
<thead>
<tr>
<th>Roles</th>
<th>Administrator</th>
<th>Parent</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Rating</td>
<td>2.81 (A)</td>
<td>2.82 (A)</td>
<td>3.10</td>
</tr>
</tbody>
</table>

Roles sharing a common line are not statistically different at the .05 level.

Discussion

Results from this study expand current knowledge about the teacher-selection process from a procedural as well as a substantive perspective. Procedurally, most legislative studies within the professional literature have been explored largely from a compliance perspective relative to uniform federal guidelines governing the employee selection process. Within these type of studies, employment policies (ADEA, ADA, or Title VII) have been held constant, and specific characteristics of applicants (age, ethnic group, disability status, or sex) have been varied to assess
Effects of a State Mandated Policy

the compliance of school administrators relative to these legislative acts. In our study, we did the opposite.

That is, we were able to vary employment policies addressing a specific phase of the teacher-selection process while holding constant characteristics of the teacher candidate. Through using a quasi control group (see Campbell & Stanley, 1963) unavailable for research focusing on uniform federal mandates that are constant across states, we are able to assess effects of rather than compliance with a specific employment policy (KERA) initiative formulated by a legislated body. As a result of this different slant for exploring teacher selection decisions, our findings indicated that this particular policy initiative involving the use of site-based teacher councils has little effect on outcomes at the screening stage of the teacher-selection process (see Table 2).

Indeed, teacher screening decisions emanating from a state mandating the use of site-based teacher councils (Kentucky) were found to be similar to those occurring in a state relegating the teacher-selection process to school administrators (Ohio). Although there is likely some degree of blurring with respect to actual teacher selection practices within each state, our manipulation in this area is not weak from a legislative perspective. From a legislative perspective, one state has mandated the use of site-based decision making relative to teacher selection (our experimental group, Kentucky), while the other state (our control group, Ohio) has failed to decentralize through legislation the teacher-selection process.

Although it might be hypothesized that many districts use a collaborative process for selecting teachers (including our control state), such a hypothesis is questionable when it comes to screening as opposed to selecting (interviewing) teachers. Within our control state (Ohio), some urban school districts (e.g., Akron City School District) use only administrators for screening teachers to insure that affirmative actions goals are met, other suburban school districts (e.g., Dublin City Schools) use commercial teacher screening instruments (e.g., Teacher Perceiver Instrument) administered only by corporate trained administrators for screening teacher applicants, and many rural school districts screen teacher applicants at the administrative level before involving additional stakeholders (teachers and/or parents) in the interviewing process because interviewing is time intensive and costly (Cable & Gilovich, 1998). Within our experimental state involving SBCs, the process is different than in our control state.

Further strengthening our manipulation of this legislative initiative involving the decentralization of the teacher-selection process is a temporal factor. In each state, the procedure used to screen and select teachers has been operating for over a decade. As such, each state should be well practiced in its particular legislated method for selecting teachers.

Potentially confounding findings relative to vesting selection with either a legislated decentralized process versus a legislated centralized teacher-selection process are the actual roles of participants taking part in the selection process. No doubt, some potential roles may be more capable than other potential roles when it comes to the selection of teachers, and existing teacher selection research has examined teacher selection almost exclusively from the perspectives of school administrators. Others have noted, however, that perspectives vary by role of the evaluator and that individuals lacking specific role experience in a particular focal position use different information for evaluating potential job candidates than those possessing role experience (Barr & Hitt, 1987).

Within the context of our study involving KERA, the roles deemed most important are parents, principals, and teachers because these constituents are viewed as those closest to students and best equipped to make teacher selection decisions (Appalachia Educational Lab, 1992; Knapel, Moore, Coe, & Aagaard, 1995; Lindle & Shrock, 1993). To assess the potential confounding involving the teacher-selection process (decentralized versus centralized) and the role of the evaluators (parents, principals, or teachers) within this process, we manipulated the latter as well as the former potential sources of systematic variance within the teacher screening process through
using a fixed-effect experimental design. Although we failed to detect either a main effect or an interaction effect involving legislated action on the part of state mandates involving the teacher-selection process (see Table 2), we did find a statistically significant main effect for role of a potential screening member within the teacher-screening process (see Table 2).

Teachers evaluate hypothetical teacher candidates consistently higher than either principals or parents on one of the dependent variables assessed in our study (see Table 4). Unlike principals or parents, teachers as members of the screening committee rated hypothetical teacher candidates higher on their professional competencies as assessed by the composite evaluation score focusing on job skills. However, with respect to offers for an interview, all role incumbents rated hypothetical teacher candidates similarly (see Table 3).

At least two explanations are offered for the above finding. One explanation is based on role theory that suggests “a person’s role determines not only behavior but also beliefs and attitudes” (Michener et al., 2004, p. 8) and to alter these attributes requires a role change on the part of the decision maker. As such, teachers screening applicants would likely view all teacher applicants favorably on skill related criteria unless there is compelling evidence to the counter because these are the same skills teachers use everyday in their current role.

The other explanation is based on the political ramification associated with devaluing the job skills of a potential colleague. That is, teachers performing screening decisions may fail to devalue a job candidate’s skills within the selection process because the former may fear that an applicant could become ultimately a co-worker in their immediate job setting. To learn subsequently that a colleague rated job qualifications of an applicant low within the selection process would fail to promote future harmonious relationships among faculty in the event that an applicant becomes an employee within the same building.

To explore the utility of these different potential explanations for the main effect involving the role of the evaluators as detected within this study, future research is needed in this particular area of investigation. With respect to our first explanation that is nested within context of role theory, only teachers held a role congruent with the hypothetical applicant, and it is unclear if these findings are a function either of the specific roles manipulated (administrator, parent, or teacher) or of the role congruence/incongruence among screeners. Consequently, future research should vary the role of screeners within the selection process to include focal applicant positions located within the same building but varying to include a congruent role position (as done in this study) as well as a role incongruent position (e.g., counselor) within the same experimental context.

Likewise, to determine the validity of our explanation concerning political ramifications associated with devaluing the credentials of a potential colleague by teachers within the same school building, additional research is warranted that varies the assignment of teacher candidates (within a building versus within a different building). Other research drawing on social distance theory has shown that principals use different standards for evaluating teacher candidates depending on whether the applicant is to be assigned to the same school building (proximal evaluation) or to a different school building (distal evaluation), and teachers may well form similar types of impressions (Oto, 2003). Until these later types of investigations are conducted, it is unclear which of our explanations (role congruence or role proximity) is most viable for the current findings.

In addition to exploring the effects of policy enactments on the part of state governments (decentralized versus centralized decision making) and the role of different evaluators (parents, principals, or teachers) within the teacher screening process, we examined the outcomes of the teacher screening process for low-performing and high-performing school districts. Conventional wisdom fueled, at least in part, by a national study (Wise, et al., 1987) may have led many to assume that high-performing school districts do something different from low-performing school districts when it comes to the procurement of teacher candidates as employees. Within our study we tested
Effects of a State Mandated Policy

We found that screening decisions for hypothetical teacher candidates were the same regardless of an elementary school’s accountability level as defined by a standard set of criteria used in both states (see Table 2). This last finding, like our previous finding, is amenable also to at least two explanations both of which have implications for research and practice involving the teacher-selection process in the field setting. One explanation is that studies contrasting elementary schools defined as different relative to accountability measures may offer little insight about improving the teacher-selection process because these schools perform similarly when it comes to screening teacher candidates.

The other explanation for our finding rests with our specific experimental protocol. Within our study, qualifications of teacher candidates were held constant and personnel in both high- and low-performing elementary schools evaluated the same teacher candidate that was equally qualified (or unqualified) in all experimental conditions. In reality, low-performing schools may fail to enjoy as an adept applicant pool as high-performing schools, and low-performing schools, although equally proficient as high-performing schools in the teacher-selection process, must choose among the less able teacher candidates in the field setting. That is, selection processes/outcomes would be the same for equal candidate pools but different for unequal candidate pools.

This second explanation for our findings has some support within the professional literature. Winter and Morganthal (2002) found that an important factor influencing applicants’ attraction toward vacant focal positions is the academic achievement of assigned students, and this finding implies that given a choice among schools the most able candidates seek only high-performing schools. To cast additional light on our findings relative to selection practices in high- and low-performing schools, future research should vary both the achievement levels of schools and the quality of teacher candidates within the same study.

Not to be overlooked when considering the above mentioned speculations are the operational parameters used to define low- and high-performing schools. We used a median split based on a linear combination of composite scores (proficiency test, promotion rates, and attendance data) to define low-performing and high-performing schools. Because relative levels of performance were defined within states, this process assumes that talent is similar and equally distributed within both our experimental and control state. However, to assume otherwise implies that elementary students are better endowed in one state as compared to the other state and would have produced a statistically significant (and undetected) interaction effect with these data if academic accountability influences screening decisions.

When differentiating between high and low performance in each state among elementary schools, we used a median split. This operational decision has both disadvantages as well as advantages. The major disadvantage is that a potentially stronger operational method of differentiating between low- and high-performing schools exists. Samples could have been drawn from the extremes of the distribution on accountability measures for elementary schools (e.g., more than one standard deviation from the mean). Clearly, sampling from the extremes of the distribution would have provided a much stronger manipulation by eliminating approximately 68% of the schools in each state. As such, the basic research question would have changed from one exploring how low- and high-performing schools may differ within the selection process to one that explores how the lowest of the low and highest of the high schools may differ within the teacher-selection process.

Interestingly, even if we had chosen the stronger manipulation and even if we had found a statistically significant statistical effect for accountability, our same research question as posed in this study would remain still unanswered and would plague still policy makers from a practical
perspective in guiding their deliberations about altering selection practices. That is, what is best for most elementary schools and not just the highest 16% the lowest 16%? Although this later research question remains unanswered on the basis of our data, a statistically significant effect with our approach (that we failed to detect) would have been the more informing.

In contrast to the approach contrasting extremes, we sought to capture screening decisions typical of all schools in both states given the legislative thrust of our study. To differentiate between low- and high-performing schools in our study, we performed a median split on the basis of a composite linear score that included all schools and not a select few. This choice of an operational definition for low and high performance on an accountability measure has some specific tradeoffs as noted in the method section of this manuscript and as discussed in the earlier paragraphs.

Because we failed to detect statistically significant effects for academic accountability among elementary schools with our data, considerably more research is warranted in this area, research that approaches academic accountability of elementary schools from a truly experimental viewpoint as opposed to an applied legislative perspective as has been done in this study. There are still questions whether selection practices at the screening stage of the selection differ for the extreme low and high performers. Although we would speculate that valid selection practices would be realized more likely by considering school specific demand characteristics than by mimicking selection practices of the elite, only additional research can provide such information through comparing those school performing far from state averages.

Collectively, these results imply other specific recommendations that warrant the consideration of policy makers and educational personnel. First, changes in the status quo (centralized versus decentralized selection practices) should be based on empirical evidence rather than strictly a political agenda. If empirical information is lacking relative to alternative teacher selection practices as noted within the literature, then these governing bodies should fund site-based studies to guide their decision making process before enacting any changes on a statewide basis.

Second, attempts to mimic the teacher selection practices of high-performing elementary schools in low-performing elementary schools is without any empirical support in the general selection literature and the data in this study. Such attempts to study only high-performing schools as done in the past (Wise, et al., 1987) ignore other important contextual factors beyond the teacher that influence the performance of public schools. Based on our data, such future efforts require attention to the composition of applicant pools relative to quality, the role congruence of evaluators and applicants within the teacher-selection process, and especially to traditional selection paradigms focusing on pre-employment assessments and post-employment outcomes.

Finally, our study, like all studies, suffers from specific limitations. Most importantly, we chose specific states (Kentucky and Ohio) varying teacher-selection processes, manipulated selective role incumbents (parents, principals, and teachers) within the teacher-selection process, defined academic accountability of elementary schools according to certain criteria, focused on a particular type of focal position (elementary teacher candidate), used the same set of teacher credentials across all experimental conditions, examined only screening as opposed to interviewing decisions, and assessed screening decisions relative to specific criteria. Consequently, any generalization beyond these restraints is pending further investigations.
References


About the Authors

I. Phillip Young
University of California–Davis

Kimberly Miller-Smith
Southwest Licking County Local School District–Ohio

Email: ipyoung@ucdavis.edu

I. Phillip Young is a professor of education at the University of California –Davis and co-directs both a Joint Doctoral Program in Educational Leadership and the University of California Educational Research Center. His research interests include human resource issues in education.

Kimberly Miller-Smith is Assistant Superintendent at Southwest Licking Local School District. Her research interests are human resource issues and organizational change strategies.
Archivos Analíticos de Políticas Educativas

Associate Editors
Gustavo E. Fischman & Pablo Gentili
Arizona State University & Universidade do Estado do Rio de Janeiro

Founding Associate Editor for Spanish Language (1998—2003)
Roberto Rodríguez Gómez

Editorial Board

Hugo Aboites
Universidad Autónoma Metropolitana-Xochimilco

Dalila Andrade de Oliveira
Universidade Federal de Minas Gerais, Belo Horizonte, Brasil

Alejandro Canales
Universidad Nacional Autónoma de México

Erwin Epstein
Loyola University, Chicago, Illinois

Rollin Kent
Universidad Autónoma de Puebla, Puebla, México

Daniel C. Levy
University at Albany, SUNY, Albany, New York

María Loreto Egaña
Programa Interdisciplinario de Investigación en Educación

Grover Pango
Foro Latinoamericano de Políticas Educativas, Perú

Angel Ignacio Pérez Gómez
Universidad de Málaga

Diana Rhoten
Social Science Research Council, New York, New York

Susan Street
Centro de Investigaciones y Estudios Superiores en Antropología Social Occidente, Guadalajara, México

Antonio Teodoro
Universidade Lusófona Lisboa,

Adrián Acosta
Universidad de Guadalajara, México

Alejandra Birgin
Ministerio de Educación, Argentina

Ursula Casanova
Arizona State University, Tempe, Arizona

Mariano Fernández
Enguita Universidad de Salamanca, España

Mariano Narodowski
Universidad Torcuato Di Tella, Argentina

Vanilda Paiva
Universidade Estadual Do Rio De Janeiro, Brasil

Mónica Pini
Universidad Nacional de San Martín, Argentina

José Gimeno Sacristán
Universidad de Valencia, España

Nelly P. Stromquist
University of Southern California, Los Angeles, California

Carlos A. Torres
UCLA

Claudio Almonacid Avila
Universidad Metropolitana de Ciencias de la Educación, Chile

Teresa Bracho
Centro de Investigación y Docencia Económica-CIDE

Sigfredo Chiroque
Instituto de Pedagogía Popular, Perú

Gaudencio Frigotto
Universidade Estadual do Rio de Janeiro, Brasil

Roberto Leher
Universidade Estadual do Rio de Janeiro, Brasil

Pia Lindquist Wong
California State University, Sacramento, California

Iolanda de Oliveira
Universidade Federal Fluminense, Brasil

Miguel Pereira
Catedrático Universidad de Granada, España

Romualdo Portella do Oliveira
Universidade de São Paulo

Daniel Schugurensky
Ontario Institute for Studies in Education, Canada

Daniel Suarez
Laboratorio de Políticas Publicas-Universidad de Buenos Aires, Argentina

Jurjo Torres Santomé
Universidad de la Coruña, España