J Bus Psychol (2013) 28:401–410
DOI 10.1007/s10869-013-9299-4

# Legal Risk in Selection: An Analysis of Processes and Tools

**Kate Z. Williams · Meline M. Schaffer · Lauren E. Ellis**

Published online: 4 April 2013


**Abstract**
*Purpose* This paper reviews a decade of employment litigation to illuminate the most legally dangerous selection devices and employment practices.
*Design/Methodology/Approach* A sample ($n = 312$) of court cases drawn from 10 years of Bloomberg BNA case briefs was analyzed to determine which *selection tools* (e.g., biographical information blank, interview, cognitive ability test, and psychomotor test) and which *selection processes* (e.g., violations of the four-fifths rule, administrative inconsistencies, lack of documentation, failure to provide accommodations) are most at risk for litigation for unfair employment practices.
*Findings* Results demonstrate that while some selection tools do attract legal scrutiny, dangerous hiring practices such as favoritism against protected classes and improper human resource documentation put employers at far greater risk of suit. When considering cases settled outside of court and those that continued to trial, the data reveal that employers lose employment discrimination cases at a rate nearing 90 % and suffer an average payout of over $1.5 million per case.
*Implications* Just as legal challenges once drove the search for selection tools free of adverse impact, the current legal landscape demonstrates the necessity of fair and consistent selection processes. This paper provides evidence of common mistakes in implementing selection systems—mistakes that lead to costly legal battles.
*Originality/Value* This paper reduces cumbersome legal records into useful evidence of trends in recent employment law cases. Selection system designers and organizations who implement them will benefit from avoiding the risky hiring practices presented in this paper.

**Keywords** Protected class · Employment discrimination · Employment selection · Employment litigation · Risk

In a competitive business environment, hiring and promoting the best workers is of primary concern to organizations. To accomplish this goal, organizations seek valid selection tools that consistently identify the best possible applicant. Individuals with higher selection scores should perform more effectively on the job (Coward and Sackett 1990). These tools are not perfect, however, as some tools result in subgroup differences based on race and gender. Organizations that use unfair selection practices are held accountable by legal requirements to avoid discrimination in hiring. Therefore, organizations must balance the sometimes competing goals of hiring the best workers, establishing a diverse workforce, and avoiding costly legal challenges.

Plyburn et al. (2008) note that organizations have difficulty simultaneously maximizing validity and hiring a diverse workforce, calling this challenge the "diversity-validity dilemma." Some of the most valid selection procedures have lower pass rates for members of racial minority groups and women. The social, ethical, and business outcomes associated with employing a diverse

This paper is based on a poster presentation delivered at the Society of Industrial/Organizational Psychology 25th Annual Conference, 2010.

K. Z. Williams · M. M. Schaffer · L. E. Ellis
Department of Psychology, Clemson University, Clemson, SC, USA

K. Z. Williams (✉)
Tri-County Technical College, 7900 Hwy 76, Pendleton, SC 29670, USA
e-mail: kate.zaner.williams@gmail.com

workforce makes it undesirable to use a selection system that creates such subgroup differences, regardless of the strength of the system's validity. Conversely, sacrificing validity in favor of protecting diversity might undercut an organization's competitive advantage gained through a highly skilled workforce.

The current study adds another layer of challenge to the diversity-validity dilemma, that of legal disputes. Not only do organizations have ethical and business reasons to build a diverse workforce, but in doing so—by avoiding selection tools that create subgroup differences in hiring—the organization protects itself from costly legal challenges. Using the most valid selection tools might both undermine diversity *and* open the organization to legal risk due to discrimination in hiring. Title VII of the 1964 Civil Rights Act forbids discrimination on the basis of race, color, national origin, religion, or sex. Under the threat of litigation, firms face a difficult choice between using a valid but legally dangerous selection device or a somewhat less valid but less litigious tool.

Terpstra et al. (1999) reviewed the legal consequences of nine types of selection devices: (1) unstructured interviews, (2) structured interviews, (3) biographical information blanks, (4) cognitive ability tests, (5) personality tests, (6) honesty tests, (7) psychomotor tests, (8) work sample/performance tests, and (9) assessment centers. They examined 20 years of federal court cases through content analysis to determine the frequency and outcomes of complaints about these tools.

Terpstra et al. (1999) found unstructured interviews were the most frequently challenged selection tool accounting for 57 % of all cases. In addition to unstructured interviews, cognitive ability tests and psychomotor tests were significantly over-represented in court. Structured interviews, work samples, and assessment centers were significantly under-represented.

The outcomes of these cases were also reviewed by Terpstra et al. (1999). Overall, the defense prevailed most of the time, with cases about structured interviews most likely to survive legal scrutiny; all of the cases about structured interviews were decided in favor of the defendant. Work sample tests also performed well, with 87 % of cases decided for the defendant, followed by cognitive ability tests (67 %), unstructured interviews (59 %) and psychomotor tests (58 %). In summary, Terpstra et al. (1999) found that the unstructured interview is most likely to result in litigation and is among the least likely to result in a positive outcome for the company. Conversely, structured interviews, work samples, and assessment centers are least likely to result in a suit, with structured interviews being the most defensible in court. Similarly, other research (Williamson et al. 1997) prior to the Terpstra et al. (1999) study indicated that less structured interviews are more likely to result in litigation while those that are highly objective and job related are more likely to hold up in court.

It is important to note that differences in litigation rates are expected due to differences in the rate of use of each selection tool, applicant reactions, adverse impact of selection techniques, and validity. Selection devices are used at differential rates across organizations, meaning that even if each selection tool was equally litigious, the percentages of cases observed would trend toward those selection devices that are used most frequently by organizations. Applicant reactions can also influence the likelihood to file legal complaints. For example, applicants who perceive a violation of procedural justice, such as limited job-relevance of predictors or little opportunity to demonstrate related abilities, might be more likely to sue. The literature shows negative applicant reactions are associated with structured interviews, biographical information blanks, cognitive ability tests, personality tests, honesty tests, and psychomotor tests. Conversely, favorable applicant reactions are more strongly associated with unstructured interviews, work samples, and assessment centers (Terpstra et al. 1999). These applicant reactions are not perfectly aligned with legal defensibility but may be more associated with intent to litigate than with the outcome of the litigation.

Some types of selection devices might be more likely to withstand legal challenge because of their validity. General cognitive ability is most often accepted as offering the strongest predictive validity, with meta-analytic results ranging from $r = .48$ (Bertua et al. 2005) to $r = .62$ (Salgado and Anderson 2003). In their review of the legal defensibility of employment interviews, Williamson et al. (1997) review meta-analyses of over 100 studies that reported validities of structured interviews of $r = .24$ to $r = .34$ and validities of unstructured interviews of $r = .11$ to $r = .18$. Psychomotor tests are good predictors of performance for certain types of jobs, with validities ranging from $r = .40$ (Hunter and Hunter 1984) to $r = .53$ (McHenry et al. 1990). Moderate validity coefficients are cited for biographical data ($r = .37$, Hunter and Hunter 1984). Finally, the predictive validity of personality tests ($r = .23$, Barrick et al. 2001) and honesty/integrity tests ($r = .34$ to $r = .47$, Berry et al. 2007) might also be of interest when considering the likelihood of legal action.

More than ten years have passed since Terpstra et al.'s (1999) research. Do these findings still hold in today's legal landscape? The current paper replicates Terpstra et al.'s research by examining the selection tools and procedures addressed by employment litigation from January 1, 1998 to July 1, 2010. While Terpstra's original paper included only litigation that originated from selection devices, the way these tools are actually used by companies, or their

selection process, is a large source of legal complaints. To this end, this paper will first analyze the types of selection tools and selection processes that are more likely to be the target of legal challenge. Second, although Terpstra included only challenges to hiring decisions, the current paper will include both hiring and promotion decisions. Finally, because complaints and outcomes are not perfectly related, this paper will evaluate the outcomes associated with each type of selection tool and selection process.

## Method

Bloomberg BNA (formerly the Bureau of National Affairs) publishes information and analyses for business and government, including the Employment Discrimination Verdicts and Settlements database covering employment law cases. In Terpstra et al.'s (1999) original research, the BNA's Fair Employment Practice Cases from 1978 through 1997 were used to identify all federal court cases in the United States that involved hiring discrimination. Their study included only cases that involved hiring or selection discrimination (105 total) and ignored cases that focused on other employment decisions, such as promotion, job assignments, pay, or training.

In replicating this study, Bloomberg BNA's membership-restricted database, Employment Discrimination Verdicts and Settlements, was used to gather data. This database includes federal and state cases, as well as complaints filed by the Equal Employment Opportunity Commission (EEOC) and the United States Labor Department's Office of Federal Contract Compliance Program (OFCCP). As such, this database now offers a more complete resource of employment law complaints than was available to Terpstra et al. (1999). In fact, one of the limitations that Terpstra notes, that the federal court cases included in their study may not be representative of the larger population of legal complaints, is rectified in the current study because cases that went to trial and cases that were settled prior to court are now included. Furthermore, the current paper reports not only on hiring decisions, but promotion decisions as well.

Because the Bloomberg BNA database includes such a wide range of cases, an initial screening of the data revealed that the nine selection devices investigated by Terpstra et al. (1999) were not sufficient to code the array of hiring practices that were challenged in legal cases. Eight selection tools were identified in the current study, with those duplicating the Terpstra study indicated with an asterisk: unstructured interview*, structured interview*, other interview (unknown structure), biographical information blank*, cognitive ability test*, psychomotor test*, and other test. The new category of "other test" included tests that did not fit into the existing test categories, such as industry-specific hiring tests. Integrity, work sample, and assessment center, which were included in Terpstra et al. (1999), were coded and then dropped from further analysis because of the low frequency of occurrence in the current sample.

In reviewing these legal cases it became apparent that many of the complaints focused not on the selection tool itself, but on the way selection tools were used, or the employers' *selection processes*. For this reason, we further extended Terpstra et al.'s (1999) work by coding cases based on violations in selection process. The following selection process complaints were identified and will be defined in detail: use of problematic criteria, violations of the Four-fifths Rule, administrative inconsistencies, personal bias or favoritism, lack of documentation, quota or unlawful affirmative action programs, unfair recruiting source, and failure to provide accommodations.

Problematic criteria included cases where the employer made hiring or promotion decisions without pre-established criteria, used shifting criteria, used non-job-related criteria, or used criteria that unlawfully discriminated against members of the protected classes. For example, in *EEOC v. Mike Fink Corporation* (2010), the defendant argued that their company policy to hire only male servers was necessary to preserve the historic accuracy of their restaurant's theme. This males-only criteria was challenged in this case. In another example (*Port Authority Police Asian Jade Society v. Port Authority of New York*, 2009) it was alleged that there were no standard criteria for promotion decisions. Instead promotions seemed to be based more on personal favoritism.

The "Four-fifths Rule" or "80 % Rule" is the common procedure used by courts to estimate the statistical disparity between protected classes produced by the organization's hiring or promotion procedure (Twomey 2005). In these cases, plaintiffs argued that the organization hired minority candidates at a rate less than 80 % of the selection rate of majority applicants. While most of these cases were concentrated under OFCCP complaints, these Four-fifths Rule complaints were also brought by EEOC and private plaintiffs.

In cases cited for administrative inconsistency, some applicants received inconsistent treatment based on their membership in a protected class. For example, one defendant marked the application form of Black applicants to indicate they should not be hired for in-room care at a nursing home (*Hill v. Merrill Gardens*, 2005) while in *Brown v. Alabama Department of Transportation* (2010), the names of Black applicants would "disappear" from the list of potential hires and would "reappear" after a White applicant was hired.

Bias and personal favoritism cases include traditional disparate treatment cases where employers purposefully

treated people differently based on their protected class membership. These cases include defendants who used derogatory language, refused to hire, and blatantly made hiring decisions based on protected class. For example, in *Hartman v. Albright* (2000), the manager told the plaintiff that she would not be hired because she was a woman.

With lack of documentation, employers kept incomplete or improper records that impeded their ability to demonstrate the fairness of their selection system. For instance, in *Bumphus v. Timec* (1998), the company shredded the plaintiff's original race discrimination complaint. In *EEOC v. Griffith Rubber Mills* (1998), the defendant violated laws by comingling medical records and application forms.

Complaints about quota argued that the organization had illegally reserved a certain number of positions for candidates based on their membership in protected classes. For example, managers at one restaurant distributed a company policy via email requiring 80 % of newly hired bartenders be women and the remaining 20 % be men (*EEOC v. Razzoo's*, 2008).

Cases coded as recruiting source used recruiting practices that discouraged or prevented members of protected classes from applying, such as failing to post open positions and relying solely on word-of-mouth recruiting to obtain applicants. In *EEOC v. Phoenix Suns* (2003), the defendant printed job advertisements that included an explicit preference for male applicants to join the basketball team's "Zoo Crew" entertainment squad. In another example, *Commissioner Tucker v. Nob Hill General Stores* (1998) involved a complaint that only male employees were informed of promotional opportunities, leading to an underrepresentation of female managers within the organization.

Finally, failure to provide accommodations includes cases where employers unlawfully prohibited applicants from applying for hiring or promotion due to their need for accommodations. For example, in both *EEOC v. Americall Group* (2008) and *United States v. Baltimore City Public Schools* (2001), the defendants failed to accommodate seeing-eye dogs for visually impaired applicants.

Coding began after these categories were defined a priori by the first and second author. The Bloomberg BNA Employment Discrimination Verdicts and Settlements provides summaries of court decision, with objective data in searchable fields and details about the lawsuits in case briefs.

First, the Bloomberg BNA database was searched for all cases between 1998 and 2010 that were filed under these types of discrimination: age, gender/sex, national origin, religion, race, disability/disabilities. Gender, national origin, religion, and race are the protected classes introduced by the Civil Rights Act of 1964. Color is also protected by CRA but is not included in this study due to the low base rate of cases. Disability is protected by the Americas with Disabilities Act of 1990 while age is protected by the Age Discrimination in Employment Act of 1967.

The authors recorded the following objective data provided by the database: name and date of the case, whether the complaint was based on a hiring or promotion decision, the protected classes included in the complaint, the outcome of the case, and the payment amount (if any). The case briefs were content-analyzed to determine the selection tool (i.e., test) and selection process (i.e., recruiting source). In more than a third of the cases (37.7 %), there was insufficient information about the tools or processes in question to include the case in the study. Cases were allowed to receive more than one code within the categories of hiring/promotion, protected class (i.e., age and race), selection tool, and selection process.

The results of the first and second authors' coding were compared for agreement. On average, the two raters agreed on 44 % of the cases. Agreement meant an exact match for all possible variables, which ultimately required agreement on 27 criteria (i.e., all levels of each variable were required to match). As expected, interrater agreement was strongest for the objective searchable data fields; interrater agreement for hiring/promotion, protected class, and outcome was 91, 89, and 82 %, respectively.

Most discrepancies resulted within the content-analyzed material when a category was assigned by one but not both authors. Stated another way, because cases were permitted to receive multiple categorical assignments—especially in selection tools and processes—initial interrater agreement was low because of a failure to assign all possible categories. In the second round of coding, each reviewer reconsidered the disputed cases and one or both raters edited their codes. The second round of coding resulted in agreement on 94 % of the 508 cases. The third author reviewed the remaining 40 cases under dispute following the same coding scheme as the first two authors. This third round of coding broke the tie in 85 % of these cases, resulting in 502 cases with 100 % rater agreement and 6 cases that were eliminated from analysis for lack of agreement.

Although Terpstra et al. (1999) left out promotion cases, complaints about promotion include similar selection issues as those involved in hiring and were therefore included in this study. These promotion cases fit the framework of the study and allowed for a richer understanding of employment discrimination. Like Terpstra, cases involving pay, training, reduction in force, layoffs, seniority, or other employment actions were excluded. These cases allowed for a clearer picture of legal threat to employment selection practices, rather than exhaustive human resources practices, which are beyond the scope of this paper.

## Results

A total of 502 complaints were included in the present study. Of the 502 cases, some of which contain complaints about multiple issues and were therefore coded into multiple categories, 69 (13.7 %) cases included a complaint against a selection device, 275 (54.7 %) were associated with a problem in the selection process, and 190 (37.7 %) cases did not have sufficient information to determine the reason for the complaint. These 190 cases were eliminated from further analysis, leaving a final sample size of 312.

The potential for a single case to be coded into one or more of 6 selection tool categories and into one or more of 8 selection process categories explains much of the low interrater agreement in the first round. After full rater agreement was established, the overlapping codes within and between categories were evaluated. There were 69 cases with a selection tools complaint, of which 90 % centered around only one selection tool while the remaining 10 % included misuse of two selection tools. Of the 275 cases with a selection process complaint, most (70 %) focused on only one process, another 20 % included 2 selection processes and the remaining 10 % complained of violations of 3 or 4 selection processes. There is also overlap between complaints about at least one selection tool and at least one selection process: 34 of the 312 cases (11 %) contain both a selection tool and selection process complaint.

Protected classes are the groups established by the Civil Rights Act of 1964, the Americas with Disabilities Act of 1990, and the Age Discrimination in Employment Act of 1967. The largest groups of cases included discrimination based on gender (33 %) and race (32.5 %), followed by disability (11.9 %), age (11.4 %), national origin (8.3 %) and religion (2.8 %).

Table 1 lists the percentage of cases filed against each selection tool in the present study and also includes numbers from the Terpstra study for comparison. While unstructured interviews drew over half (58 %) of the complaints during the 25 years (1978–1997) summarized in the Terpstra study, the last decade found only 4 % of cases based on unstructured interviews. As expected, unstructured interviews produced a greater occurrence of litigation than structured interviews (3.9 % vs. 1.3 %), but at such a low base rate as to be essentially inconsequential. Interviews with unknown level of structure comprised 13.2 % of the selection tool cases. Unfortunately, the data do not include enough information to determine what characteristics of these interviews were challenged.

Although unstructured interviews were the most challenged selection tool in the original study, selection tests were the largest type of selection tool under fire (54.7 %) in the current study. The largest category of selection test was "other," which prompted the authors to recode these 24 cases into test subcategories. A total of 13 cases were moved into a new "fire, police, and civil service exam" category, while the remaining 11 cases were maintained in the "other test" category because information about the specific test was not available (7 cases), or because the number of cases in the sub-category were too small to be meaningful (one case requiring a test of English proficiency, one case requiring a driving test, two cases requiring a written mechanical aptitude test).

**Table 1** Legal challenges for discriminatory selection tools

| | Terpstra et. al. (1999) | | Present study | |
|---|---|---|---|---|
| | *n* | % | *n* | % |
| Biographical information Blanks | 0 | 0.0 | 13 | 17.1 |
| Interview (structured) | 9 | 5.6 | 1 | 1.3 |
| Interview (unstructured) | 91 | 57.6 | 3 | 3.9 |
| Interview (unknown) | | | 10 | 13.2 |
| Test (cognitive ability) | 28 | 17.7 | 5 | 6.6 |
| Test (personality) | | | 1 | 1.3 |
| Test (physical) | 22 | 13.9 | 17 | 22.4 |
| Test (other) | 8 | 5.1 | 26 | 34.2 |
| Total | 158 | | 76 | |

In addition to pre-employment tests, application forms or Biographical Information Blanks (BIBs) made up 20 % of the challenges to selection tools. These cases generally centered around the inclusion of illegal questions at the pre-employment stage, such as medical, mental health, and disability inquiries prior to hiring. For example, in *EEOC v. AMR Eagle* (2000), the company included a long list of diseases on their pre-employment screening application that effectively eliminated applicants with disabilities from flight attendant jobs.

A new contribution of this study is the analysis of the legal challenges to selection processes, including the use of problematic criteria, violations of the 4/5 rule, administrative inconsistencies, personal bias or favoritism, lack of documentation, quota or unlawful affirmative action programs, unfair recruiting source, and failure to provide accommodations. Table 2 describes the results.

As demonstrated in Table 2, blatantly biased selection processes are most likely (32 %) to lead to a legal challenge. Companies should not be surprised when they are sued for maintaining organizational preferences for applicants from specific protected classes or for employing decision-makers who make derogatory comments about applicants and employees. Complaints about inconsistent hiring processes (22 % of cases) demonstrate the importance of maintaining a consistent selection process. Unlawful use of a quota system seems to draw the fewest

**Table 2** Legal challenges for discriminatory selection processes

| | *n* | % |
|---|---|---|
| Criteria | 64 | 16.6 |
| 4/5 rule | 42 | 10.9 |
| Inconsistent | 84 | 21.8 |
| Bias/favoritism | 122 | 31.6 |
| Lack of documentation | 17 | 4.4 |
| Quota | 11 | 2.8 |
| Recruiting | 23 | 6.0 |
| Accommodation | 23 | 6.0 |
| Total | 386 | |

complaints (3 %), as might be expected due to the strict guidelines governing the use of court-ordered quotas to correct past discrimination.

## Cases That were Settled Versus Decided in Court

Of the total 312 cases with determinable causes, 82 cases settled by the involved parties prior to trial, 75 cases settled out of court by the EEOC and 22 complaints settled out of court by the OFCCP for a total of 179 cases (57 %) settled prior to trial. A total of 133 cases (43 %) were tried at the state, district or federal court level.

Cases that went to trial were compared to cases that were settled prior to trial (Table 3). Of the 75 complaints about a selection instrument, just over half (51 %) were settled before trial. Interestingly, a full 75 % of cases involving complaints about interviews—structured or unstructured or unknown—were settled out of court.

Switching attention now to cases that surrounded selection processes, over half (57 %) were settled prior to trial. An overwhelming majority of cases involving violations of the 4/5 rule, recruiting source, and accommodations were settled prior to trial (81, 91, and 87 %, respectively). Because the OFCCP inspects companies with federal contracts, when violations of the 4/5 rule are uncovered, employers seem to prefer the efficiency of settling prior to trial. It may be that these cases (4/5 rule, recruiting source, accommodations) are better supported by factual data than the other selection process cases. For example, investigations launched by the EEOC often occur because the company's hiring data demonstrates statistical violations of the 4/5 rule. Because the 4/5 rule is based on the number of minorities hired compared to those in the population, the recruiting sources would also come under fire in these investigations. The data gathered for these cases may make OFCCP and EEOC cases stronger than cases brought by non-government parties. In fact, nearly all (97 %) of the 103 complaints brought as a result of investigations by OFCCP and EEOC resulted in findings against the company, with 92 % of these cases settled outside of court. By comparison, fewer cases brought by non-government plaintiffs (39 %) were decided by settlement.

## Outcomes of Cases at Trial

When companies refuse to settle and proceed to trial they win only 26 % of cases. Organizations lose cases about their selection tools at nearly the same rate as they lose cases about their selection processes (78 % to 73 %, respectively). Taking a broader view of the data, a full 89 % of both the selection tool cases and the selection process cases were found to be discriminatory, which includes those settled outside court and cases that went to trial and resulted in decisions against the employer.

## Monetary Awards

When deciding to continue to trial or to settle prior to court, employers must weigh the costs and benefits of each decision. Prior to analysis, the monetary awards were screened for outliers. Two cases received extremely large awards ($508 million and $253 million) and were removed from further analyses. Across the remaining 250 cases in which the plaintiff received an award, the average award was $2,161,974. After examining cases that were more than 3 standard deviations away from this mean, three more cases ($47 million, $46 million, and $20 million) were identified and eliminated from analysis, with the final mean award of $1,727,222 for the remaining 247 cases.

Cases surrounding a selection tool received an average award of $1,617,061, while cases that centered on a selection process received an average award of $2,064,963, making it more costly for employers to be sued over a selection process than a selection tool. The total number of selection tools challenged in each case and the total number of selection processes challenged in each case were used to create two new interval variables, which were then correlated with the award amounts. There was a small but significant positive correlation between the number of selection processes challenged in each case and the size of the award amount ($r = .10$, $p < .05$) while there was no correlation between the number of selection tools challenged in each case and the size of the award amount for those cases.

Settling out of court resulted in average fees per case of $802,228 for EEOC cases and $683,353 for OFCCP cases. As discussed previously, because the vast majority of cases brought by the EEOC and OFCCP are found in favor of the plaintiff, organizations may be encouraged to reach a faster and cheaper solution to these complaints by settling with

**Table 3** Comparison of cases tried in court to cases settled out of court

| | Total | Settled | Tried in court | Plaintiff prevailed | Defendant prevailed |
|---|---|---|---|---|---|
| | *n* | *n* | *n* | *n* | *n* |
| Selection tools | | | | | |
| BIB | 13 | 9 | 4 | 3 | 1 |
| Interview (structured) | 1 | 0 | 1 | 1 | 0 |
| Interview (unstructured) | 3 | 0 | 3 | 1 | 2 |
| Interview (unknown) | 10 | 3 | 7 | 7 | 0 |
| Test (cognitive ability) | 5 | 4 | 1 | 0 | 1 |
| Test (psychomotor) | 17 | 9 | 8 | 6 | 2 |
| Test (other) | 26 | 13 | 13 | 11 | 2 |
| Tools total | 75 | 38 | 37 | 29 | 8 |
| Selection processes | | | | | |
| Criteria | 64 | 37 | 27 | 19 | 8 |
| 4/5 rule | 42 | 34 | 8 | 6 | 2 |
| Inconsistent | 84 | 40 | 44 | 37 | 7 |
| Bias/favoritism | 122 | 53 | 69 | 49 | 20 |
| Documentation | 17 | 11 | 6 | 5 | 1 |
| Quota | 11 | 5 | 6 | 3 | 3 |
| Recruiting | 23 | 21 | 2 | 1 | 1 |
| Accommodation | 23 | 20 | 3 | 1 | 2 |
| Processes total | 386 | 221 | 165 | 121 | 44 |
| Grand total | 461 | 259 | 202 | 150 | 52 |

the regulatory associations rather than risk trial. Cases that were settled out of court by individual plaintiff rather than by the EEOC/OFCCP averaged $3,529,908. In comparison, cases that went to trial and were found in favor of the plaintiff cost organizations an average of $1,341,985. It may be the case that many of the cases settled prior to trial were for more egregious violations which could be reflected in the larger monetary awards for these cases. It is likely that defense attorneys counsel their clients to settle when faced with well-documented, egregious complaints of discrimination, but to proceed to trial for those cases in which the plaintiff has a weak or undocumented case.

It should be noted that not all cases received million dollar settlements. Approximately a quarter of the plaintiffs (20 % of those who settled; 27 % of those who went to trial) received payouts of less than $100,000. At the other end of the spectrum, nearly a third of the plaintiffs (30 % of those who settled; 31 % of those who went to trial) received monetary awards of one million dollars or more.

## Discussion

The data provided here is compelling evidence supporting the importance of careful consideration of the legal defensibility of selection tools and procedures. The likelihood of being sued, in addition to the typical outcomes of such complaints, can help employers effectively plan selection tools and selection programs that are relatively safe from litigation. Next, specific outcomes of interest and recommendations for use are provided.

### Selection Tools

Because interviewing is one of the most commonly used selection tools, these cases are of particular interest. Most of the interview cases (71 %) did not contain enough information to determine the structure of the interview in question. Previous research found that structured interviews—those with greater objective job relatedness, more standardized administration, and with multiple interviewers—are more likely to hold up in court (Williamson et al. 1997). More information may be needed before a conclusion can be drawn about the nature of the interviews that attract legal scrutiny.

Regarding employment testing, the data clearly show the importance of using job-related instruments. On average, employers prevailed in only 10 % of the cases addressing their employment tests, including both cognitive ability and psychomotor tests. For example, in *EEOC v. NationsBank of Tennessee* (2001) a cognitive ability test discriminated against Hispanic employees by requiring English proficiency that was not necessary for the job. Psychomotor tests also had serious content validity problems. For example, in *EEOC v. Consumers Energy* (2001), the company used a pre-employment strengths test for meter readers and janitor positions that demonstrated adverse impact for gender. Smartly, the case was settled prior to trial because the test measured skills that were not required for these positions. When psychomotor tests did go to trial, the defendant prevailed in only 25 % of the cases. This data make it clear that with both cognitive ability and psychomotor tests, the test must be job related to be legally defensible.

A decrease in the number of legal challenges was observed for several selection tools. In Terpstra et al. (1999), unstructured interviews comprised 58 % of the cases while this same selection tool accounted for only 4 % of the suits in the current data. Likewise, integrity tests, work samples, and assessment centers were included in Terpstra et al. (1999), but these assessment tools appeared in the current data in such small numbers that they were eliminated from analysis. Although speculative, base rate

usage, applicant reactions, and validity of the tools might explain the decrease in legal challenges.

A decrease in the base rate of usage of these instruments would decrease the possibility of legal action against those tools. The literature is clear about the inferior validity of unstructured interviews compared to structured interviews (Ulrich and Trumbo 1965; Hunter and Hunter 1984), which could be driving down the overall usage of unstructured interviews between the years included in Terpstra et al. (1999) and the present study. According to Levy (2010), integrity tests are administered to over 5 million applicants each year, while at least 50 % of major companies report using assessment centers, which suggests this explanation may not be adequate for those tools.

Application reactions might also explain differential litigation levels. Selection tools with strong face validity might foster applicants' feelings of procedural justice, decreasing their likelihood to file suit. This proposal is in keeping with the application of procedural justice theories to the selection context (Gilliland 1994), which suggests that applicants will judge the selection process as fair if the tests appear to be job related, if the applicant is provided with an opportunity to express his abilities, if the questions appear appropriate, and if the applicant is extended positive interpersonal treatment. Applicant reactions can only explain the decrease in litigation, however, if applicant reaction to these tools increased in the time between Terpstra et al. (1999) and the current study. It is conceivable that applicant reaction could have increased if the base rate of these tools (integrity tests, work samples and assessment centers) increased and applicants are more familiar with these procedures. The data in the present study offers no information to defend this hypothesis; further investigation is needed.

Finally, tools with stronger validity could avoid suit, either because applicants accept the assessment as fair based on perceptions of face validity or because attorneys fail to pursue cases that are unlikely to success (i.e., those focused on tools with strong validity). In support of this explanation, validity studies demonstrate strong coefficients for integrity tests ($r = .34$; Berry et al. 2007), work samples ($r = .39$; Roth et al. 2005), and assessment centers ($r = .37$; Gaugler et al. 1987). However, this explanation would not apply to unstructured interviews due to the weak validity previously discussed.

## Selection Processes

This study makes a unique contribution to the literature by including an analysis of the way selection processes affect litigation. The greatest number of suits included biased hiring decisions (32 % of process suits), followed closely by inconsistent administrative practices (22 %) and poorly defined criteria (17 %). As examples of inconsistent processes, the US Court of Appeals described one hospital's selection process as "a peculiarly informal process" because their explanations for not hiring the plaintiff were different from the written job description, giving the decision "a flavor of post hoc rationalizations" (*Dennis v. Columbia Colleton Medical Center*, 2002). In *Dunlap v. Tennessee Valley Authority* (2008), the court ruled the company's hiring process was discriminatory because they found 70 counts of manipulating test scores and changing the ratio of interview and test scores in candidate rankings.

Although recruiting practices comprise a relatively small set of process cases (6 %), nearly all these cases were settled prior to trial. In *Allen v. Tobacco Superstore* (2007), the company relied on word-of-mouth to publicize open positions and had no consistent procedures for advancement; employees simply asked a supervisor to be considered. The court found the word-of-mouth hiring and promotion process—which resulted in a company-wide dearth of Black store managers despite operating in communities with large Black populations—was discriminatory.

The documentation available to the plaintiff may explain the variance in the rate at which selection process complaints are settled. As described in the results section, cases centered on the 4/5 rule, recruiting source, and accommodations were settled at a greater rate than other selection process cases. Organizations would be more motivated to settle with a plaintiff who can provide evidence of discrimination. A single plaintiff who was not hired or promoted is likely to have minimal access to company data they can present as proof of discrimination. Even blatant instructions to "not hire any more old plumbers" (Potence v. Hazleton Area School District, 2004) would be difficult for an individual plaintiff to prove. As such, plaintiffs with complaints surrounding personal favoritism, biased criteria, and documentation might not be able to offer a fully documented case. The EEOC and OFCCP, on the other hand, have access to historical data of hiring trends and have the power to launch a stronger fact-gathering investigation. These statistically-based cases comprise the 4/5 rule category and would also be expected to appear in the recruiting source and documentation categories.

## Double Jeopardy: Combining Tools and Processes

By content-analyzing the available data, this research attempted to tease apart the differential effects of selection tools and processes to shed light on which tools and practices emerge as best practices to withstand legal challenges. When applied in an organization setting, however, these tools and processes rarely operate

independently. It is common for an organization to use a combination of selection tools within a selection system—using both a structured interview and a psychomotor test, for example. These tools can then be unit- or regression-weighted and combined into a composite score (Cascio 1998).

If tests with unequal validity are unit-weighted, this can have the effect of washing out the higher validity (i.e., combining an interview with average validity of .14 with an assessment center score with average validity of .54 would produce an average validity of the selection system that is lower than the validity of the single best predictor).

Practitioners make decisions about the utility of the overall selection system based on a number of factors, including the diversity-validity dilemma. The legal defensibility of the tools is another characteristic that should be considered. Increasing validity through the inclusion of a legally-risky selection tool might not increase the overall utility of the selection system.

### Limitations and Future Research

The current study is only as complete as the historical information on which it is based. A total of 190 cases did not contain sufficient information to be properly coded for inclusion in this study. The lack of information makes it impossible to tell whether these cases differed in some substantial way from the ones described here. In addition, a few of the categories, such as cognitive ability tests with 5 cases, contained such low base rates that conclusions could not be drawn here. Finally, there was not enough information provided about the cases in general to determine what meaningful differences there were between cases in which the defendant prevailed or cases that were settled or lost during trial. It seems likely that organizations with stronger validity studies to support their practices or hiring tests that were based on thorough job analysis would be more likely to prevail in court. Unfortunately, the data provided did not allow us to examine these questions.

The current paper uncovered the shift of risk from selection tools to selection processes in the past 10 years. While these results comprise an important contribution of this paper, they also highlight the need for continued monitoring of the legal landscape for future changes in employment litigation. While selection devices are often the source of litigation, this study shows that the legality of selection procedures is of even greater importance. The value of these findings lies in highlighting the comparative danger of *how* selection procedures are used versus the instruments themselves. In fact, over 83 % of employment discrimination cases focused on complaints on employers' selection processes rather than selection tools. Once an employment discrimination case is filed, employers stand a limited chance of a successful outcome. With only 26 % of all cases decided in favor of the employer, organizations' best strategy is to avoid employment litigation.

To avoid litigation, employers should clearly define and document their selection process and then follow that process, insuring each candidate has a fair and equal opportunity to apply. This includes insuring that jobs and promotions are fairly advertised to diverse groups in the first place. Employers must also insure that they can demonstrate job relatedness of employee tests and selection criteria and that policies explicitly banning employees based on gender (i.e., *EEOC v. Mike Fink Corporation*, 2010) will not be tolerated. Companies should ferret out managers who make biased hiring decisions, if not for ethical reasons at least to protect themselves from the most common cause of employment discrimination law suits.

Given the results of this study, employers can make a well-informed decision about the competing needs of their selection programs: to be legally safe and psychometrically sound. While these findings are not causal and will not guarantee a positive legal outcome, knowing more about the comparative level of legal risk associated with each tool and process will help employers successfully negotiate the complex legal landscape of selection.

## References


Allen v. Tobacco Superstore (2007). 475 F.3d 931 (8th Cir. 2007).

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection & Assessment, 9*(1/2), 9–30.

Berry, C. M., Sackett, P. R., & Wiemann, S. (2007). A review of recent developments in integrity test research. *Personnel Psychology, 60*(2), 271–301.

Bertua, C., Anderson, N., & Salgado, J. F. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. *Journal of Occupational and Organizational Psychology, 78*(3), 387–409.

Brown v. Alabama Dep't of Transp (2010). 11th Cir., No. 08-14371.

Bumphus v. Timec (1998). N.D. Cal., No. C-96-03585 SI.

Cascio, W. F. (1998). *Applied Psychology in Human Resource Management* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Commissioner Tucker v. Nob Hill General Stores (1998). EEOC, Charge No. 370-94-0117.

Coward, W. M., & Sackett, P. R. (1990). Linearity of ability-performance relationships: A reconfirmation. *Journal of Applied Psychology, 75*, 297–300.

Dennis v. Columbia Colleton Medical Center (2009). 290 F.3d 639 (4th Cir. 2002).

Dunlap v. Tennessee Valley Authority (2008). 206 F.8a 121 (6th Cir. 2008).

EEOC v. Americall Group (2008). N.D. Ill., No. 04-C-5554.

EEOC v. AMR Eagle Inc. (2000). N.D. Tex., No. 3:98-CV-0763-M

EEOC v. Consumers Energy (2001). No. 98-70846.

EEOC v. Griffith Rubber Mills (1998). N.D. Ind., No. 98-CV-0183.

EEOC v. Mike Fink Corporation (2010). M.D. Tenn., No. 3-96-0790.

EEOC v. NationsBank of Tennessee (2001). E.D. Tenn., No. 3: 00-CV-170.

EEOC v. Phoenix Suns (2003). D. Ariz., No. CIV 02-0963 PHX JAT.

EEOC v. Razzoo's (2008). N.D. Tex., No. 3:05-cv-0562-P.

Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology, 72*, 493–511.

Gilliland, S. W. (1994). Effects of procedural and distributive justice on reactions to a selection system. *Journal of Applied Psychology, 79*, 691–701.

Hartman v. Albright (2000). No. CA 77-2019 JR.

Hill v. Merrill Gardens (2005). N.D. Ind., No. 1:04-cv-00248-TLS-RBC.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*(1), 72–98.

Levy, P. E. (2010). *Industrial Organizational Psychology: Understanding the organization* (3rd ed.). New York, NY: Wadsworth.

McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project a validity results: The relationship between predictor and criterion domains. *Personnel Psychology, 43*, 336–354.

Plyburn, K. M., Ployhart, R. E., & Kravitz, D. A. (2008). The diversity-validity dilemma: overview and legal context. *Personnel Psychology, 61*, 143–151.

Port Authority Asian Jade Society v. Port Authority of New York and New Jersey (2010). 05 Civ. 3835.

Potence v. Hazleton Area School District (2004). No. 03-1535, 03-2647.

Roth, P. L., Bobko, P., & McFarland, L. A. (2005). A meta-analysis of work sample test validity: Updating and integrating some classic literature. *Personnel Psychology, 58*(4), 1009–1037.

Salgado, J. F., & Anderson, N. (2003). Validity generalization of GMA tests across countries in the European community. *European Journal of Work and Organizational Psychology, 12*(1), 1–17.

Terpstra, D. E., Mohamed, A. A., & Kethley, R. B. (1999). An analysis of federal court cases involving nine selection devices. *International Journal of Selection and Assessment, 7*(1), 26–34.

Twomey, D. P. (2005). *Employment discrimination law: A manager's guide* (6th ed.). Mason, OH:Thomson/South-Western.

Ulrich, L., & Trumbo, D. (1965). The selection interview since 1949. *Psychological Bulletin, 63*(2), 100–116.

United States v. Baltimore City Public Schools (2001). D. Md., No. 01 CV 4187.

Williamson, L. G., Campion, J. E., Malos, S. B., Roehling, M. V., & Campion, M. A. (1997). Employment interview on trial: Linking interview structure with litigation outcomes. *Journal of Applied Psychology, 82*, 900–912.