Kolbe Corp EEOC Information

Legal Compliance

A. Legal Principles for Employment & Advancement

The most important case impacting employee testing is <u>Griggs v. Duke Power</u>, 401 US 424 (1971), in which the United States Supreme Court held that the employer, Duke Power, had established unlawful racially discriminatory criteria for employment and advancement, including testing and educational requirements. The legal principles emerging from that case and its legal progeny are:

> 1)   Any testing or other system of selection or classification, even if facially neutral, which has a "disparate impact" on a protected group including religion, national origin, age, gender, and handicap, will be the basis for an action for employment discrimination, unless the employer can demonstrate that there is a substantial "business necessity" for the practice. See <u>James v. Stockham Valves and Fittings Co.</u>, 559 F.2d 310 (5th Cir. 1977), cert. den. 434 U.S. 1034 (1978) (evidence of disparate impact of employment test on Blacks combined with gross disparities between numbers of Whites and Blacks in positions requiring high scores sufficient to establish adverse impact). However, the impact must be shown in the specific job setting. See <u>Adams v. Texas & Pacific Motor Transport Co.</u>, 408 F. Supp. 156 (E.D. La. 1975), (employer's use of the same test found inappropriate in <u>Griggs</u> was not held to be facially invalid without a showing of disparate impact among employer's job applicants). Disparate impact has been defined by EEOC Guidelines to constitute, "A selection rate for any race, sex, or ethnic group which is less than four-fifths(4/5)(or eighty percent) of the rate for the group with the highest rate." 41 C.F.R. section 60-3.4(d). However, this "four-fifths rule" has been criticized by some courts. See <u>Clady v. County of Los Angeles</u>, 770 F.2d 1421 (9th Cir. 1985).
>
> 2)   In order to satisfy the "business necessity" requirement, a test or requirement must be proven to be "job related." See <u>Brunet v. City of Columbus</u>, 642 F. Supp. 1214 (S.D. Ohio 1986), (employer established job-relatedness of mechanical reasoning test); and <u>United States v. LULAC</u>, 793 F.2d 636 (5th Cir. 1986) (lower court erred in not considering the job-relatedness of a biased test).
>
> 3)   Tests used for employment selection and classification can not be validated as "job related" in the abstract, but the inferences which are drawn from the test results which are used by an employer in

employment decisions may only be validated within a specific employment context. See Albemarle Paper Co. v. Moody, 422 U.S. 405 (1975) (citing EEOC Guidelines, "Discriminatory tests are impermissible unless shown, by professionally acceptable methods, to be 'predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated.' 29 CFR section 1607.4(c).")

4)   A facially discriminatory pattern can be established by showing a statistically significant difference between the hiring patterns of an employer and an appropriate reference group (usually adult population in the geographical area). See International Brotherhood of Teamsters v. United States, 431 U.S. 324 (1977). However, courts have sometimes allowed general population statistics to be used to prove disparate impact. See Dothard v. Rawlinson, 433 U.S. 321 (1977) (national statistics on height and weight used to show disparate impact on females of Alabama requirements for prison guards)

5)   Once a facially discriminatory pattern is established, the employer has the burden to establish the "business necessity" or "job relatedness" of the test or other selection procedure. Additionally, the employer must show that the employment practice which selects or classifies, even if it measures a "job-related" "business necessity" does so in a way which is less discriminatory than other available alternatives and does not inappropriately use scores of those near the top of a line of progression to exclude applicants without considering the effect of work experience and on-the-job training. See Albemarle Paper Co. v. Moody, supra, (use of tests proscribed where they had not been validated for all jobs for which they were used and where scores of experienced workers were used to establish cut scores without considering on-the-job development).

B.   Government Guidelines, Standards & Case Law

The government has provided guidelines for appropriate methods for statistical validation of employment practices, in the EEOC's "Uniform Guidelines on Employee Testing Procedures" (1978), (Guidelines) which incorporate the "Standards for Educational and Psychological Testing," (Standards) jointly published by the American Educational Research Association., The American Psychological Association. and the National Council on Measurement in Education. Although the Guidelines are not law, they have been given deference by the courts. See Ensley Branch, NAACP v. Seibels, 14 F.E.P. Cases 670 (N.D. Ala. 1977)(courts should "follow the guidelines adopted by the EEOC...absent some cogent reason" for not doing so).

The Guidelines, Standards and case law have favored criterion-based employment testing, which typically infers probable standing in measure of job performance (criterion) from the applicant's test score. See Douglas v. Hampton, 512 F.2d 976 (D.C. Cir. 1975)("courts that have expressed a view on the relative merits of these techniques have uniformly manifested a preference for empirical [criterion-based] validity").   The validation study must be conducted for a specific job by the employer. See Albemarle, supra.   The stronger the relationship between the test scores and job performance and the larger the number of job duties correlated with the score, the greater the likelihood that the validation study will be acceptable. 41 C.F.R. section 60-3.14(B)(6). However, some courts have found other acceptable methods of validation. See Washington v. Davis, 426 U.S. 229 (1976) (non-criterion-based validation found adequate to establish job-relatedness). The present Guidelines accept content validation studies, involving evaluations of an applicant's skill such as typing tests or weight lifting requirements, and construct-based testing which involves testing more general traits, such as "intelligence" or "affability" which may have some relationship to the job requirements. Content testing is only appropriate where there is "evidence that the selection procedure measures and is a representative sample of the knowledge, skill, or ability" necessary to perform the job. 41 C.F.R. section 60-3.15(c)(4) see Smith v. Olin Chemical Corp., 555 F.2d 1283 (5th Cir. 1977)(job criteria such as a "good back" for a laborer may be so obvious as to preclude the need to formally establish business necessity). Construct testing requires identifying both job requirements and the personal characteristics "believed to underlie successful performance of these critical or important work behaviors." 41 C.F.R. section 60-3.14(D)(2). Such testing must further involve criteria-based studies to show that the construct is related to job performance, 41 C.F.R. section 60-3.3(B), and consequently offers no particular advantages in comparison to the other acceptable validation techniques.

As noted above, whatever testing approach is used, it may not use "vague, subjective" supervisor's rankings, nor may it measure employees at entry on the same criteria as those near the top of a line of progression without detailed consideration of factors such as promotion, speed, and on-the-job training. See Albemarle, supra.

Section 15(A) of the Guidelines establishes the record-keeping requirements for validation studies and selection testing. All validation studies and selection systems are required to track applicants by protected status to ensure that the selection technique is either bias free, or, that failing, that it is appropriately validated for the job in question and is less biased than other alternatives. 41 C.F.R. section 60-3.15(A). If there is a correlation between a biased test and objective performance (a job-related work necessity) the test may be used, but there is an on-going obligation to review other testing techniques which are less

biased. See <u>Albemarle</u>, supra. When possible, the courts will try to identify the specific portion of an employment practice gives rise to minority bias and for which employers are required to keep records that differentiate between the elements of a selection procedure. 41 C.F.R. section 60-3.15(A)(2)(a)-(b).

The Civil Rights Act of 1991 was enacted, in part, to legislatively overturn the decisions in <u>Patterson v. Mclean Credit Union</u>, 491 U.S. 164 (1989) which limited the scope of Section 1981 relief to hiring and promotion, and <u>Wards Cove Packing Co., Inc. v. Antonio</u>, 490 U.S. 642 (1989), and <u>Texas Dept. of Community Affairs v. Burdine</u>, 450 U.S. 248 (1981), which shifted the burden of persuasion to the plaintiff through the entire course of a discrimination case. In order to refute a prima facie case of employment discrimination under <u>Wards Cove</u>, the employer was simply required to articulate a business justification for its action. The 1991 act eliminates the <u>Patterson</u> distinction between hiring and other employment-related decisions, reinforces and extends the burdens for the employer in "disparate impact" cases by reducing the standards of proof necessary to file and prove a case to pre-<u>Wards Cove</u> standards, and increases the burdens of employers in testing by proscribing "race-norming" of employment tests, a practice by which employers would establish different scoring systems or cut scores for minorities.

C. Kolbe and Federal Law

1.   Compliance Studies

Consistent with the requirement for job-specific validation, the Kolbe A™ index is a non-subjective criterion-based test whose criteria may be correlated with job-specific criteria (such as sales productivity, absenteeism, etc.) through correlating scores with objective criteria reflecting job-related criteria.

The Kolbe A index is free from bias both in general and as a selector for specific jobs. Consistent with the Civil Rights Act of 1991, there have never been separate norms for any groups. Although the law does not allow evidence of the lack of bias of an instrument generally to obviate the need for evidence of lack of bias in a specific job, the Kolbe selection system ensures that, consistent with EEOC Guidelines, the Kolbe will not select any race, national origin, gender, age, or disability less than 80% as frequently as the most frequently selected group.

Initial Kolbe Study

Dr. Robert T. Keim of the Decision Systems Research Center of Arizona State University conducted an extensive study of bias of the Kolbe A index in 1990, in which he examined 4030 index results which were grouped into 17 groups reflecting common conative patterns similar to job selection criteria. Dr. Keim initially performed analyses of variance with each of the four conative instincts as dependent variables and race, gender and age as independent variables. In 65 of 68 analyses of variance the results showed that at the .05 level of significance that differences in scores on the Kolbe A index could not be attributed to the dependent variables of race, gender or age. For the three values where the initial analysis of variance did not provide conclusive results, a Chi Square analysis was conducted by computing a Chi Square base-model value for each with gender, race and age. Subsequent analyses of variance and Chi Square values were computed leaving out each of the independent variables. Comparisons between the base-model values and the subsequent values demonstrated that in no case do the independent variables of race, gender or age explain differences in scores. Dr. Keim concluded, "We can conclude that at the Alpha=.05 level the Kolbe is not biased by gender, age or race."

General Selection Study

In a subsequent selection-bias study performed in 1992, 24,416 Kolbe A index results were studied. The index results were cross-tabulated by each of 51 professions and 10 professional levels. In each profession and level in which there was an adequate minority sample (30 or more) the data was analyzed to determine whether the Kolbe A index would have selected any minority group (determined by the federally protected categories of race, gender and age) less than 80% as frequently as the most frequently selected group (the criteria for adverse impact established by the EEOC). In no category in which there was an adequate minority sample would the index have adversely selected on minority status. There was no evidence that the Kolbe A index would have an adverse impact on any minority group if used as part of a properly designed selection process.

National Origin Study

When a mode-by-mode distribution of insistence, accommodation and prevention is compared for respondents of US and non-US origin, the results show there are no statistically significant differences. The study included 10,124 respondents of US origin and 1,182 of non-US origin. The charts below reflect the percentage of respondents in each mode by zone

of intensity. Country of origin clearly does not influence the distribution of Kolbe A index results.

United States Origin

	Insist	Accommodate	Prevent
FF	39.29%	51.10%	9.61%
FT	20.89%	55.81%	23.30%
QS	36.08%	35.40%	28.52%
IM	6.87%	49.03%	44.09%
Total	25.79%	47.83%	26.38%

Non-United States Origin

	Insist	Accommodate	Prevent
FF	40.61%	50.00%	9.39%
FT	26.40%	52.37%	21.24%
QS	32.66%	34.35%	32.99%
IM	6.60%	48.39%	45.01%
Total	26.57%	46.28%	27.16%

1. **STATISTICAL STUDIES**

Selection Case Study

A later study for a Fortune 500 company selecting entry-level employees, the researcher established suggested cut scores then monitored applicants' scores to determine whether the instrument would result in any racial group or gender bias by any group being selected less than 80 percent as often as the most frequently selected groups. To determine selection the acceptable scores were pooled and the relationship to the most frequently selected was used to determine whether there was any group selected less than 80% as frequently as the most frequently selected group.

FF by Race

	Count Exp Val Residual	White 1	Black 2	Hispanic 3	Asian 4	Row Total
			RACE			
Prevent	1	4 3.2 .8	0 .8 -.8	1 .5 .5	0 .6 -.6	5 3.3%
Accommodate	2	46 45.9 .1	11 10.9 .1	6 6.6 -.6	9 8.5 .5	72 47.4%
Initiate	3	47 47.9 -.9	12 11.3 .7	7 6.9 .1	9 8.9 .1	75 49.3%
	# of Respondents	97	23	14	18	152
	% of Total Respondents	63.8%	15.1%	9.2%	11.8%	100.0%

Chi-Square	Value	DF	Significance
Pearson	2.32911	6	.88708
Likelihood Ratio	3.50314	6	.74355
Mantel-Haenszel test for linear association	.13126	1	.71713

Minimum Expected Frequency -.461
Cells with Expected Frequency < 5 - 4 of 12 (33.3%)
Number of Missing Observations: 0

FT by Race

	Count Exp Val Residual	White 1	Black 2	Hispanic 3	Asian 4	Row Total
			RACE			
Prevent	1	11 10.8 .2	3 2.6 .4	1 1.6 -.6	2 2.0 .0	17 11.2%
Accommodate	2	63 63.2 -.2	15 15.0 .0	11 9.1 1.9	10 11.7 -1.7	99 65.1%
Initiate	3	23 23.0 .0	5 5.4 -.4	2 3.3 -1.3	6 4.3 1.7	36 23.7%
	# of Respondents	97	23	14	18	152
	% of Total Respondents	63.8%	15.1%	9.2%	11.8%	100.0%

Chi-Square	**Value**	**DF**	**Significance**
Pearson	2.18646	6	.90178
Likelihood Ratio	2.21033	6	.89937
Mantel-Haenszel test for linear association	.14634	1	.70205

Minimum Expected Frequency -.1.566
Cells with Expected Frequency < 5 - 5 of 12 (41.7%)
Number of Missing Observations: 0

QS by Race

	Count Exp Val Residual	RACE				
		White 1	Black 2	Hispanic 3	Asian 4	Row Total
Prevent	1	36 37.7 -1.7	9 8.9 .1	7 5.4 1.6	7 7.0 .0	59 38.8%
Accommodate	2	46 46.6 -.6	11 11.0 .0	6 6.7 -.7	10 8.6 1.4	73 48.0%
Initiate	3	15 12.8 2.2	3 3.0 .0	1 1.8 -.8	1 2.4 -1.4	20 13.2%
	# of Respondents	97	23	14	18	152
	% of Total Respondents	63.8%	15.1%	9.2%	11.8%	100.0%

Chi-Square	Value	DF	Significance
Pearson	2.38997	6	.88057
Likelihood Ratio	2.62383	6	.85436
Mantel-Haenszel test for linear association	1.06619	1	.30181

Minimum Expected Frequency -.1.842
Cells with Expected Frequency < 5 - 3 of 12 (25.0%)
Number of Missing Observations: 0

IM by Race

	Count Exp Val Residual	RACE				
		White 1	Black 2	Hispanic 3	Asian 4	Row Total
Prevent	1	46 41.5 4.5	8 9.8 -1.8	4 6.0 -2.0	7 7.7 -.7	65 42.8%
Accommodate	2	49 54.2 -5.2	15 12.9 2.1	10 7.8 2.2	11 10.1 .9	85 55.9%
Initiate	3	2 1.3 .7	0 .3 -.3	0 .2 -.2	0 .2 -.2	2 1.3%
	# of Respondents	97	23	14	18	152
	% of Total Respondents	63.8%	15.1%	9.2%	11.8%	100.0%

Chi-Square	Value	DF	Significance
Pearson	4.24266	6	.64388
Likelihood Ratio	4.94943	6	.55031
Mantel-Haenszel test for linear association	.93180	1	.33439

Minimum Expected Frequency - .184
Cells with Expected Frequency < 5 - 4 of 12 (33.3%)
Number of Missing Observations: 0

FF by Gender

		Gender		
	Count Exp Val Residual	Female 1	Male 2	Row Total
Prevent	1	1 1.1 -.1	4 3.9 .1	5 3.3%
Accommodate	2	17 15.6 1.4	55 56.4 -1.4	72 47.4%
Initiate	3	15 16.3 -1.3	60 58.7 1.3	75 49.3%
	# of Respondents	33	119	152
	% of Total Respondents	21.7%	78.3%	100.0%

Chi-Square	Value	DF	Significance
Pearson	.29073	2	.86471
Likelihood Ratio	.29038	2	.86486
Mantel-Haenszel test for linear association	.17544	1	.67532

Minimum Expected Frequency – 1.086
Cells with Expected Frequency < 5 - 2 of 6 (33.3%)
Number of Missing Observations: 0

FT by Gender

	Count Exp Val Residual	Gender		
		Female 1	Male 2	Row Total
Prevent	1	4 3.7 .3	13 13.3 -.3	17 11.2%
Accommodate	2	19 21.5 -2.5	80 77.5 2.5	99 65.1%
Initiate	3	10 7.8 2.2	26 28.2 -2.2	36 23.7%
	# of Respondents	33	119	152
	% of Total Respondents	21.7%	78.3%	100.0%

Chi-Square	Value	DF	Significance
Pearson	1.18224	2	.55371
Likelihood Ratio	1.14716	2	.56351
Mantel-Haenszel test for linear association	.40588	1	.52407

Minimum Expected Frequency – 3.691
Cells with Expected Frequency < 5 - 1 of 6 (16.7%)
Number of Missing Observations: 0

QS by Gender

	Count Exp Val Residual	Gender		Row Total
		Female 1	Male 2	
Prevent	1	15 12.8 2.2	44 46.2 -2.2	59 38.8%
Accommodate	2	12 15.8 -3.8	61 57.2 3.8	73 48.0%
Initiate	3	6 4.3 1.7	14 15.7 -1.7	20 13.2%
	# of Respondents	33	119	152
	% of Total Respondents	21.7%	78.3%	100.0%

Chi-Square	Value	DF	Significance
Pearson	2.48095	2	.28925
Likelihood Ratio	2.48227	2	.28906
Mantel-Haenszel test for linear association	.02406	1	.87674

Minimum Expected Frequency – 4.342
Cells with Expected Frequency < 5 - 1 of 6 (16.7%)
Number of Missing Observations: 0

IM by Gender

	Count Exp Val Residual	Gender Female 1	Gender Male 2	Row Total
Prevent	1	15 14.1 .9	50 50.9 -.9	65 42.8%
Accommodate	2	18 18.5 -.5	67 66.5 .5	85 55.9%
Initiate	3	0 .4 -.4	2 1.6 .4	2 1.3%
	# of Respondents	33	119	152
	% of Total Respondents	21.7%	78.3%	100.0%

Chi-Square	Value	DF	Significance
Pearson	.64028	2	.72605
Likelihood Ratio	1.06374	2	.58750
Mantel-Haenszel test for linear association	.24996	1	.61710

Minimum Expected Frequency – .434
Cells with Expected Frequency < 5 - 2 of 6 (33.3%)
Number of Missing Observations: 0

The study established that there were no significant differences in the cell frequencies for the cut score cells for the respective racial and gender groups. Consequently the Kolbe A index would be unlikely with this population to select any group 80% less frequently than any other group.