

**Statistical Report on the
Kolbe Indexes[®]
Dr. Ryan Thomas
February 18, 1998**

I. Theory

The Kolbe system is based on the theory of conation which premises all human behavior on the interaction between the cognitive, (knowledge) the affective, (feeling or belief) and the conative (instinct or will)¹. Named for its developer, Kathy Kolbe, the system categorizes observable behavior patterns into four categories which she identified as instincts to 1) probe, 2) pattern, 3) innovate, and 4) demonstrate. Her initial instrument was refined through the use of criterion-group analysis and correlation studies to eliminate cognitive or affective variables.

II. Test Description

The Kolbe ATM Index (Kolbe) is a forced-choice instrument which requires subjects to choose one of four responses reflecting how they would be most and least likely respond to 36 single-sentence problem-solving or behavioral scenarios. The raw scores are translated into a set of four scales which reflect the subject's conative instincts to act in terms of the tendency to initiate, respond or prevent action in each of the conative modes.

The Kolbe's use in selection is based on the hypothesis that the most successful employees in similar jobs tend to have conative insistencies which fall within a well defined range and that these ranges, in turn, tend to be consistent with the conative expectations of the job identified by supervisors and cohorts.

III. Research Report

This report includes summaries and meta-analyses of research performed by independent researchers and consultants as well as studies commissioned by Kolbe. In some instances I had access to original data, while in others I only had the statistical summaries. Where possible, I have included original statistical analyses. If there are nay particular studies where the original analysis would be helpful, I will contact the researchers to determine their availability.

The meta-analyses are comprised of similar studies that each had objective performance measurements. I have included mostly meta-analyses because they tend to reflect more accurately the viability of the test across multiple selection environments.

The first group are predictive studies in which the Kolbe was administered and subsequent performance measures were tracked. The second group are descriptive studies in which high performers were objectively identified and then the Kolbe was administered to establish a group profile. Finally, I have included illustrative bias studies and an illustrative reliability study.

A. Predictive Studies

Study One

The first study is a meta-analysis of studies conducted for the systems and computers division of Eastman Chemical and disbursements, materials control and sales teams from Hershey during 1995-1996, to predict goal attainment and profitability of employee groups based on Kolbe predictions of group viability. As I had access to the original data for these studies, and since the measurement criteria were identical, I simply aggregated the groups to perform the meta-analysis.

CORRELATION OF PROFITABILITY, GOAL ATTAINMENT & VIABILITY

Variable	Cases	Mean	Std Dev
PROFITABILITY	201	80.2189	18.9471
GOAL ATTAINMENT	201	61.4179	15.9334
VIABILITY	201	72.3184	15.4990

Variables	Cases	Cross-Prod Dev	Var-Covar
PROFITABILITY			
GOAL ATTAINMENT	201	29306.6119	146.5331
PROFITABILITY			
VIABILITY	201	53875.9900	269.3800
GOAL ATTAINMENT			
VIABILITY	201	39076.2537	195.3813

Correlations:	PROFITABILITY	GOAL ATTAINMENT	VIABILITY
PROFITABILITY	1.0000	.4854**	.9173**
GOAL ATTAINMENT	.4854**	1.0000	.7912**

VIABILITY .9173** .7912** 1.0000

N of cases: 201 1-tailed Signif: * - .01 ** - .001

The study reflects high correlations between viability and both goal attainment and profitability significant at the .001 level.

Study Two

The next study was completed by Dr. Richard S. Deems in January of 1992. He developed a Kolbe cut score for a national financial services company based on the profiles of successful managers. All 483 branch manager trainees hired in 1991 were divided into three approximately equal groups: 1) a control group which was not given the Kolbe, 2) a study group of trainees given the Kolbe whose scores fell outside the recommended range but whose managers were trained in conation to respond to the conative dissonance, and 3) a study group of trainees whose scores fell within the recommended range. At the end of six months, 11.7% of the control group had left the company for job-related reasons, 5.5% of those who were conatively mismatched, but whose managers tried to mitigate the conative dissonance by using the trainee's Kolbe results had left for job-related reasons, and none of the conatively-matched trainees left for job-related reasons.

Test Grouping/Attrition	Control	Partial use of Kolbe	Kolbe Utilized
Left for Job-related Reason	11.7%	5.5 %	0.0%
Still Employed After Six Months	88.3%	94.5%	100.0%
Selection Rate	88.3%	94.5%	100.0%

Chi-square: 19.997843 (df = 2 Chi-square at .005 = 10.5966)

Deems concluded that selection within the recommended Kolbe range eliminated job-related attrition. The Chi-square statistic corroborates that there is a statistically

significant difference between the groups using the Kolbe and the control group at the .005 level.

B. Descriptive Studies

This meta-analysis uses the combined results of eight studies which were used to identify the conative profiles of objectively measured high performers in different fields (aerospace, sales, construction, insurance, etc.) The analysis employs the Winer Combined Test².

Study	n	r	t	M ₁ -M ₂	Signif.
A	425	.86	34.668	.60	>.01
B	30	.90	10.925	.84	>.01
C	45	.69	6.216	.71	>.01
D	39	.95	18.502	.64	>.01
E	87	.93	23.325	.67	>.01
F	39	.81	8.402	.75	>.01
G	177	.68	12.269	.68	>.01
H	50	.58	4.934	.68	>.01

$$\text{Winer } Z_c = 41.35$$

$$\text{Cohen ES} = .67$$

Since the studies were descriptive, high correlations were not unexpected, nor is the Z_c of 41.35 all of which are significant at .001 level. However, of perhaps greater interest is the measure of effect size, which compares the observed correlations with established population values. Using Cohen's equation for effect size where d = the effect size index for t-tests of means in standard units, M_1 and M_2 = population means in original measurement units and s = the standard deviation of either population³, and assuming the Kolbe's use as a selection instrument, the effect size would be .67. Although no absolute standards exist for effect size, anything over .5 is considered significant⁴.

IV. Legal Compliance

Passage of the Civil Rights Act of 1991, which eased the plaintiff's burden of proof in a job discrimination case and proscribed the common practice of race-norming selection

instruments, has heightened concerns about test validity and minority bias. To comply with the law, an employer must establish that their employment practices are 1) unbiased because they create no "disparate impact" upon a protected minority group or 2) that their biased practice(s) are good predictors of job success, and no less discriminatory options are available.

Lack of Bias Study I

Dr. Robert T. Keim of the Decision Systems Research Center of Arizona State University conducted an extensive study of bias of the Kolbe instrument in 1990, in which he examined 4030 Kolbe 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 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 $\text{Alpha}=.05$ level the Kolbe is not biased by gender, age or race."

In a subsequent selection-bias study performed in 1992 by Kolbe, 24,416 Kolbe results were studied. The Kolbe 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 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 Kolbe have adversely selected on minority status. There was no evidence that the Kolbe would have an adverse impact on any minority group if used as part of a properly designed selection process.

Lack of Bias Study II

The next study was completed in January of 1994, by Dr. Clyde Stutts who administered the Kolbe to all sales employees of two branches of a nationwide

department store. The study showed that using the Kolbe for selection does not violate EEOC guidelines requiring that all suspect categories be selected by the selection criteria at least 80% as frequently as whites. Stutts established suggested cut scores for applicants consisting of an acceptable range within each Kolbe Action Mode then monitored applicants' scores to determine using the instrument resulted in any racial group or gender being selected less than 80 percent as often as the most frequently selected groups. His units of comparison were the conative zones "Resist, Accommodate, and Insist, and the cut scores were: Fact Finder (FF): Accommodate or Resist, Follow Thru (FT) Accommodate, Quick Start (QS) Accommodate or Insist, and Implementor (IM) Resist or Accommodate. 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.

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. The result for racial groups was that the least frequently selected group (Hispanics) was selected at least 91.8% as often as the most frequently selected group (Asians). The notable result is that the residual values (the difference between the expected frequency of selection or non-selection for each group and the actual frequencies) are quite small. The size of the study is important to consider when analyzing the results. For example, Asians were the most frequently selected group (77.8%), but the number of Asians selected was just .64 of a person more than expected.

SELECTION BY RACE

	RACE				Row Total
	White 1	Black 2	Hispanic 3	Asian 4	
SELECTED					
Actual %	76.3%	73.9%	71.4%	77.8%	
Actual #	74	17	10	14	115
Expected #	72.06	17.09	10.4	13.36	112.91
Residual	1.94	-.095	-.4	.64	
NON-SELECTED					
Actual %	23.7%	26.1	28.6	22.2	
Actual #	23	6	4	4	37
Expected #	24.94	5.91	3.6	4.64	39.09
Residual	-1.94	.095	.4	-.64	
# of Respondents	97	23	14	18	152
% of Total Respondents	63.8%	15.1%	9.2%	11.8%	100.0%

	Chi-Square	Value	Significance
χ^2		.2867	At the .05 level
t		12.718	At the .05 level

RESULTS BY GENDER

FACT FINDER

	Count	Gender		Row Total
		Female 1	Male 2	
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

FOLLOW THRU

	Count	Gender		Row Total
		Female 1	Male 2	
Prevent	1	4 3.7 .3	13 13.3 -.3	17 11.2%
Accommodate	2	19	80	99

		21.5	77.5	65.1%
		-2.5	2.5	
Initiate	3	10	26	36
		7.8	28.2	23.7%
		2.2	-2.2	
	# 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			

QUICK START

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

IMPLEMENTOR

		Gender	
--	--	--------	--

	Count Exp Val Residual	Female 1	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 would be unlikely with this population to select any group 80% less frequently than any other group.

V. Reliability

In a study completed in 1993, 125 individuals, representing employees of a Big Six accounting firm, and a national marketing, management and economic development firm were given the Kolbe twice with the retests occurring between eight and fifteen months from the time of the original testing. The results were analyzed using Paired T-tests and Pearson Product Moment Correlations for each conative instinct. Frequency tables were further analyzed to determine the absolute changes.

The Paired T-tests resulted in values which ranged from .87 to .50, none of which reflect a statistically significant result in the scores at the .05 level between the initial test results and the retest results. The Pearson Product Moment Correlation coefficients ranged between .67 and .88 all of which were statistically significant correlations at the .05 level for the n's studied. The analysis of the frequency tables indicated that 94.4% of all changes were within the standard error of measurement of the test of one unit.

1. Kolbe, K. (1993) *Pure Instinct*. New York: Times Books.

2. Winer, B.J. (1971) *Statistical Principles in Experimental Design* (2nd ed.) New York: McGraw-Hill.

3. Cohen, J. (1977) *Statistical Power Analysis for the Behavioral Sciences* (rev. ed.) New York: Academic Press.

4. *Supra*