Appendix B
Reliability and Validity of the Data

Two basic kinds of tests are applied to data to determine their soundness. One kind of test measures *reliability*, and the other test measures *validity*. This appendix is a small sampling of the tests that have been applied to our data.

Reliability

Reliability refers to the consistency or dependability of a measure, and you can test for reliability in a number of ways.

One method of testing for reliability is called *internal consistency*. It is the extent to which questions that are designed to measure the same attitude – say, attitude toward the work itself – correlate with each other. (This is similar to an ability test, where the questions designed to measure mathematical ability should, if they are reliable, correlate with each other.) If the questions are reliable, they should not only correlate highly with each other, but significantly more so than with questions that measure other attitude areas (just as two mathematics-ability questions should correlate more highly with each other than with two verbal-ability questions).

In most of our surveys, we intentionally ask more than one question about the same subject area. This is not to "trick" the respondents, but to get at the area from a number of angles. When we do this, we invariably find that these questions correlate highly with each other, significantly more so than they correlate with other questions on the questionnaire. Here's a small, typical example of what we mean: Table B-1 shows two pairs of questions from a survey; the two questions in each pair are designed to measure the same attitude area, more or less. The attitude areas are the attitude toward the work itself and the attitude toward the immediate supervisor. Table B.1 contains the correlations among all four questions. Notice that the correlations within each of the pairs are all high (in the .60 to .70 coefficients range) – demonstrating high reliability – and, as expected, considerably higher than the questions in one pair are with questions in the other pair (coefficients in the .40 range).

Table B-1
Correlations of Job and Immediate Manager Items

	Supervisor Is Honest	Trust my Supervisor	Sense of Accomplishment on the Job
Trust Supervisor	(.78)		
Sense of Accomplishment on the Job	.44	.45	
Use of Skills on the Job	.45	.44	(.66)

The reason that the questions across the two attitude areas still correlate significantly with each other – albeit lower than within the attitude areas – is that they are not completely independent. For example, a supervisor who employees trust is also likely to give them latitude to use their skills and abilities on the job (that might be one reason employees trust him). When more independent attitude areas are related, we find low correlations. For example, the correlations between attitudes toward the work itself and attitudes toward pay fall between .10 and .25, those between pay and equipment are .05 to .15, and so on.

Of course, we never find perfect correlations, even between questions that are designed to measure the same attitude area. For one thing, the questions tap somewhat different aspects of that area. Second, there is always "measurement error," such as an employee misinterpreting a question or making a mistake in the marking of her answer.

We once made a mistake in designing a questionnaire. We inadvertently asked two identical questions on the same very long questionnaire, once at the beginning of the questionnaire and once toward the end. This mistake turned out to be fortuitous in that it provides us with a terrific additional test of reliability, namely, "test-retest" reliability, which measures whether individuals provide consistent responses when they answer the very same questions after a short time interval. The correlation between the two was .92, which is just about as high as correlations get. This clearly demonstrates that people do not answer questions randomly or haphazardly.

Validity

Validity is a measure of the extent to which the questionnaire measures what it purports to measure. For example, attitudes toward pay might be consistent (reliable), but it might not be actually measuring attitudes toward pay (just as a thermometer might always give the same reading, but it might be the wrong reading).

We have innumerable tests of the validity of our data. Here's a sampling:

1. **Demographics**. Let's not start with attitude questions, but with what are termed demographics. In almost all of our surveys, we ask respondents for their length of service, sex, race, and so on. We do this so we can divide the data by these variables and determine whether any attitudinal differences exist between the groups (such as between men and women). These are actually very sensitive questions to ask in an "anonymous" survey because employees often feel that their answers to these questions can identify them. But, in every case where we have achieved a high response rate (more than 80%), the distributions we obtain on the surveys are almost identical to the actual demographic distributions

provided to us by the organization. Apparently, the employees, despite their misgivings about identification, tell us the truth. Table B-2 shows an example from one company.

Table B-2
Demographic Distributions: Survey Responses Versus Actual

	Percentage of Respondents in Survey	Actual Percentage in Workforce
Classification		
Full-time	92%	91%
Part-time	6%	7%
Per diem	2%	2%
Sex		
Male	53%	55%
Female	47%	45%
Tenure		
0-2 years	23%	21%
2-5 years	11%	10%
5-10 years	21%	23%
10-20 years	28%	27%
More than 20 years	18%	19%

2. Response to change. We usually survey extremely large populations. Individuals might have all sorts of quirks – an employee may, for example, be happy with his pay when, by all objective standards, his pay is low – but these quirks should not make a big difference when surveying large numbers of employees. If pay questions are valid, they should, on average, reflect actual compensation in the way compensation is normally evaluated (for example, what employees are paid compared to what they might earn in similar jobs in other companies). Therefore, if pay is significantly raised for a workforce and the answers to our questions on pay do not, on average, show increased satisfaction, we would doubt the validity of those questions. But, the fact is that when employee pay is significantly raised, their attitude toward pay *does* move up markedly.

   For example, in a company with five plants, the employees in one plant ("E") expressed much less satisfaction with their pay on a survey than did the employees in the other four plants. The reason was that the company did not differentiate in its pay scales between areas of the country and "E" was in a much higher wage and cost-of-living area than were the other four plants. This policy of no "area differentials" was changed as a result of the survey, which led to an average wage increase of about 15% over the next two years for "E" employees. In the other plants, the wages increased an average of 7%. Table B-3 shows the satisfaction with pay data from the five plants in two time periods.

Table B-3
Percent Satisfied with Pay in Five Plants in Two Time Periods

Plant	Time 1	Time 2	Difference (Time 2 - Time 1)
A	45%	46%	+1
B	53%	54%	+1
C	53%	51%	-2
D	43%	46%	+3
E	32%	49%	+17

Satisfaction with pay increased markedly in Plant E and only marginally, if at all, in the other four plants. You might ask, "What is so surprising – what would you expect?" That is precisely the point: These common-sense tests – nothing esoteric, nothing psychological – applied to attitude results give us such great confidence in their validity.

The pay data bring us to a broader point: Our surveys are performed so that the information is used; they are not for academic interest. When, for whatever reason, no action occurs after a survey, *no changes in employee attitudes are seen in the next survey*. The data are remarkably stable because it takes action to change them. Furthermore, if action is taken in one respect, it shows up on the questions relating to that change, but not in other respects. Table B-4 shows the attitudes toward the work itself in the same five plants. These attitudes hardly changed at all.

Table B-4
Percent Satisfied with the Work Itself in Five Plants in Two Time Periods

Plant	Time 1	Time 2	Difference (Time 2 - Time 1)
A	68%	67%	-1
B	74%	74%	0
C	73%	72%	-1
D	66%	69%	+3
E	76%	77%	+1

Employees carefully differentiate between the various attitude areas: Plant E markedly increased on attitudes regarding pay, but not at all regarding the work itself. This phenomenon of differentiation underlies the previously mentioned small correlation between attitudes toward pay and attitudes toward the job itself. Most employees fill out the questionnaires thoughtfully and carefully, noting what they're pleased with and what they're not.

Take another example of the impact of management action on survey data: An energy company laid off a large number of employees in two consecutive years, even though it promised after the first layoff that there would be no more. We conducted surveys after the first layoff and after the second layoff. We have no data about attitudes before the first layoff, but the results of our first survey indicated, as would be expected, that employees were unfavorable about their job security (30 percent below our norm). After the second layoff, the results, which were already very low, got even worse: now 50 percent below the norm. Table B-5 shows the data.

Table B-5
Satisfaction with Job Security

Time 1 (After First Layoff)	Time 2 (After Second Layoff)	Norm
29% favorable	9% favorable	59% favorable

Job security was the largest declining item in the second survey. Interestingly, with the loss of so many people and the subsequent demands placed on those who remained, several items showed sharp *improvements*, especially job challenge and intra-department teamwork. This makes sense

because the work became more challenging and the employees had to rally around each other to accomplish what needed to be done. Again, we see evidence of the sharp differentiation employees on surveys make between different aspects of work situations.

Take a third example of the impact of change as a test of the validity of the survey data: A particularly poor CEO (by everyone's standards, such as the board of directors, his direct reports, and the general workforce) was removed and replaced by a CEO strongly committed to improving performance and raising morale. Table B-6 shows the before and after survey results regarding the company's leadership after just one year with the new CEO at the helm.

Table B-6

Impact of New CEO on Attitudes Toward Company Leadership (% Favorable)

	Under Previous CEO	**Under New CEO**	**Difference (1 Year Later)**
Leadership Question			
Sense of direction	44%	86%	+42%
Effective leadership	42%	80%	+38%
Listening to employees	37%	72%	+35%
Overall job being done by senior management	50%	84%	+34%
Interest in employee well-being	52%	81%	+29%

As you can see, the attitude changes were huge.

3. Relationships with performance measures. Among the most important tests of validity is the way employee attitudes predict performance (just as many ability measures are supposed to predict to academic achievement.) Chapter 2, "Employee Enthusiasm and Business Success," details the strong positive relationships we and others find between employee morale and organization performance. We can provide numerous other examples of the relationship between attitudes of various kinds and performance. Here is just one more: the relationship we almost invariably find between the perceived customer focus of an organization (measured by employee surveys) and customer satisfaction (measured independently by customer-satisfaction surveys). We consistently find that, across various industries, the more employees say a company is customer-focused, the greater customer satisfaction is. Again, you might ask, "What else would you expect?" Again, that is precisely the point: What common sense tells us should be found is found, thus reinforcing confidence in the validity of the survey. No, we didn't need a survey to tell most managers how critical it is for employees to see their organizations as customer-focused to achieve customer satisfaction. The survey serves to inform management just how customer-focused its organization is.

Table B-7 shows the correlation between customer focus and customer satisfaction in companies in various industries.

Table B-7

Correlations Between Perceived Customer Focus and Customer Satisfaction

Customer Focus Question	**Type of Company**	**Correlation**
The company strives for excellence in customer service	Retail	.57
Is committed to customer satisfaction	Telecommunication	.63
Makes exceeding customer expectations a priority	Hospitality	.64
Has a strong interest in patient well being	Hospital	.56
Places management emphasis on measuring customer satisfaction	Automotive	.68