



The PDi Survey Effectiveness Index© Construct Development

by Robert A. Page, Jr., Ph.D., and Edward Tamson, Ph.D.

Contact PDi about these publications

(copyright 2002 by PDi. All rights reserved. Please do not modify, copy or distribute this article -- see website [terms and conditions of use.](#))

Executive Summary

Most organizations believe that surveys will give them the data and information measurements necessary to focus improvement efforts that serve to increase both employee and customer loyalty, thus, generate the kind of products and services which create a very positive financial ROI with surveys. For such organizations, PDi has developed a powerful measurement tool -The PDi Survey Effectiveness Index ©. This survey is valid and reliable, and provides a comprehensive assessment of important performance indicators with national comparative data on how organizational survey efforts compare with those of other companies. Incorporating the latest research on survey effectiveness, the PDi Survey Effectiveness Index © provides a measurement of the key drivers to high quality data collection and information by assessing the following essentials:

- 1. Needs Analysis**
- 2. Sampling Strategy**
- 3. Survey Design & Administration**
- 4. Survey Analysis & Feedback**
- 5. Action Planning**
- 6. Survey Synchronization**
- 7. Overall Survey Satisfaction**

The procedures used to establish reliability and validity (face, content and construct) of this index are outlined in this paper, and well as the pretest findings. Results indicate that most respondents are not satisfied with the survey efforts at their organizations - the category mean of 3.15 indicates they are neutral in their responses to positively worded questions. Survey satisfaction tends to improve with effective action planning and follow through on survey results. Respondents also tend to link survey satisfaction with survey efforts that effectively identify, analyze, and communicate important information and trends. Poor survey design and reporting issues not only tend to undermine survey satisfaction, they also tend to undermine respondent trust in management guarantees of confidentiality. Poor survey quality and credibility issues may explain the written comments noting poor respondent motivation and participation in survey efforts.

Pretest Sample.

In the fall of 2003 managers and employees working in over 50 different companies were contacted to participate in constructing a benchmark database for this Index. The list of organizations includes:

- | | |
|----------------------------------|--------------------------------------|
| 1. Ansonia Copper & Brass | 29. NEC America |
| 2. Anthem Blue Cross/Blue Shield | 30. Nicolas Moving LLC |
| 3. BAE Systems | 31. People's Bank |
| 4. Barclay's Capital | 32. Phantoms Production |
| 5. Bryan Edwards Euro-Clean | 33. Praxis Institute |
| 6. Bridgeport Hospital | 34. PriceWaterhouseCoopers LLP |
| 7. Cablevision | 35. Prudential Securities |
| 8. Computer Innovations | 36. PSEG |
| 9. Connecticut Post | 37. Red Robin |
| 10. Costco | 38. Ro-La-Lume Awning |
| 11. Council of Foundations | 39. Sabian |
| 12. CT Tax Fraud Bureau | 40. SBC SNET |
| 13. Dee's Cleaners and Laundry | 41. Schler/Lesser Architects |
| 14. Dellacamora Co. | 42. Shop Rite |
| 15. Dingere | 43. Sikorsky Aircraft |
| 16. Flame Productions | 44. Southern Connecticut State Univ. |
| 17. Fleet Bank | 45. Stanadyne Corp. |
| 18. George Harte Nissan | 46. State of Connecticut |
| 19. Hamilton Connections | 47. TC Distribution Co. |
| 20. Health Net, Inc. | 48. Trader Joe's |
| 21. Holt, Wexter, and Farnham | 49. Trefz Co. / McDonald's |
| 22. Kennedy Center | 50. Trilegiant |
| 23. Kompa Magazine | 51. U.S. Postal Service |
| 24. MacDermid Inc. | 52. Vinny's Home & Garden |
| 25. Metlife | 53. Walmart |
| 26. Merrill Lynch | 54. World Gym |
| 27. Milani Bakery | 55. Yale University |
| 28. Millburn Ridgefield Corp. | 56. Yale New Haven Hospital |

The Survey Effectiveness Index© is populating its database using NAIC codes, which replaced SIC (Standard Industry Classification) codes in 1997. For the complete listing of 2002 codes, go to: <http://www.census.gov/epcd/naics02/naicod02.txt>

Industry Cluster A: Natural Resources

- (1) Resources - agriculture, forestry, fishing, hunting, mining, oil and gas [codes 11-21]
- (2) Utilities - electric, gas, water, sewer, steam [code 22]

Industry Cluster B: Heavy Manufacturing

- (3) Construction [code 23]
- (4) Metal manufacturing - machinery, tools, equipment, transportation vehicles (land, sea, air and outer space), etc. [codes 331-333, 336]
- (5) Chemical manufacturing - including drugs, chemicals, rubber and petroleum-based products, etc. [codes 324-326]

Industry Cluster C: General Manufacturing

- (6) Food and beverage manufacturing [code 311]
- (7) Light manufacturing - textiles, apparel, wood, paper, furniture and other non-metallic products; printing services [codes 313-323, 327, 337-339]
- (8) Computer and electronics manufacturing [codes 334-335]

Industry Cluster D: Sales

- (9) Wholesale trade - durable and non-durable goods [code 423-424]
- (10) Retail trade - durable and non-durable goods [codes 44-45]

Industry Cluster E: Information Services

- (11) Communication - publishing, broadcasting, telecommunications [code 51]
- (12) Finance, insurance and real estate [codes 52-53]
- (13) Professional services - consultants, lawyers, accountants, etc. [codes 54-55]

Industry Cluster F: Private Sector Services

- (14) Transportation and warehousing - passenger services, freight shipment and storage on land, sea or air [codes 48-49]
- (15) Support services for business - temp agencies, cleaning services, collection agencies, call centers, and other service contractors [code 56]
- (16) Educational services [code 61]
- (17) Health care and social assistance [code 62]
- (18) Travel and food services - hotels, restaurants, caterers, museums, entertainment, theatres, sports and amusement parks, casinos, etc. [codes 71-72]
- (19) Personal services - home & auto repair and service, social orgs., professional associations, unions, barber shops, etc. [code 55, 81]

Industry Cluster G: Public Sector Services

- (16) Educational services [code 61]
- (17) Health care and social assistance [code 62]
- (20) Public administration - local, state and federal government [code 92]

Respondents were limited to working professionals, and were asked to choose a survey effort they were familiar with, both in terms of taking the survey, and learning of the results. Most respondents rated employee surveys, and were in the information and private sector industry clusters:

Industry Sector

	Frequency	Percent	Cumulative Percent
Valid Natural Resources	4	4.9	4.9
Heavy Manufacturing	7	8.5	13.4
General Manufacturing	2	2.4	15.9
Retailers & Wholesalers	9	11.0	26.8
Information Services	28	34.1	61.0
Private Sector Services	26	31.7	92.7
Public Sector Services	6	7.3	100.0
Total	82	100.0	

Type of Survey

	Frequency	Percent	Cumulative Percent
Valid Employee	58	70.7	79.5
Customer	10	12.2	93.2
360 Degree Profile Feedback	5	6.1	100.0
Total	73	89.0	
Missing	9	11.0	
Total	82	100.0	

Validity and Reliability

The PDi Survey Effectiveness Index© has been extensively pre-tested and meets all generally accepted statistical standards for reliability, and for face, content, and construct validity.

The PDi Survey Effectiveness Assessment Index© has been developed from reliable, valid, pre-tested questions from the extensive PDi library of questions. The index is reliable to the extent that the measures are accurate, dependable, stable, and consistent. The index is valid to the extent that the survey questions actually measure what they claim to measure. Validity addresses the basic questions of "Is this survey any good?" and "Can I trust what the findings are telling me?" Specific

types of validity include (1) face validity, (2) content validity, (3) normal distribution, and (4) construct validity.

1. Face validity assesses where the language, phrasing, and content of the survey item is clear, unambiguous, unbiased, and easily understood. Face validity problems make the survey data impossible to interpret.
2. Content validity assesses whether the survey adequately explores all relevant aspects of performance, so that the resultant data set will be comprehensive, maximizing systematic variance. Content validity problems mean the survey findings are missing or confusing important parts of the story.
3. Normal Distribution assesses whether the data for each question is normally distributed around the mean. If means are highly skewed, or the variance is restricted, the question has two problems. First, the question can not be used for higher level statistical testing, and second, you can predict the answer to these questions before you ask them again. They are not providing useful data.
4. Construct validity assesses whether the survey is performing as it should. Construct validation requires that the questions and categories feature discriminant validity (they are statistically distinct and independent from one another) and convergent validity (they correlate with each other in patterns which make conceptual sense).

Face and content validity were established by an expert panel composed of academic researchers, consultants and organizational practitioners. Once any problems identified by the expert panel were corrected, the index was pre-tested.

Normal distribution around the mean was measured through coefficients of skewness and kurtosis. Complying with generally accepted statistical standards, all questions featured skewness or kurtosis coefficients between +2.0 and -2.0.

Construct validity is assessed through measuring convergent and discriminant validity. Discriminant validity assesses whether the questions load into independent categories (called factors), or whether they are really measuring the same concept, and can be collapsed or eliminated. To test discriminant validity, the more rigorous exploratory factor analysis procedure was used. This involved a principal components analysis using an orthogonal (VARIMAX) rotation. Following generally accepted statistical standards, factors are retained if they had Eigen values greater than 1.0. Questions are retained with they featured factor loadings greater than .4, or if double-loadings are separated by a differential field greater than .10. While all of the questions successfully loaded on a factor, 5 of the 36 double-loaded. Of the 5, only one featured a sufficiently large differential field, and the other 4 were edited accordingly. They were excluded from subsequent calculations of category scores and Cronbach Alpha reliability coefficients.

Convergent validity explores if questions and categories are statistically related to each other in conceptually consistent ways. Convergent validity was explored using Pearson correlation and intercorrelation matrices. As expected, most of the categories established by factor analysis and most of the individual questions were significantly, positively related, meaning there was over a 95% probability that when the scores on one of these categories or questions were low, the other categories and questions tended to be low as well, and when scores were high, the others also tended to be high. For example, you would expect the categories of "question quality" and "survey credibility" to be positively related - you seldom see one without the other also being present.

While the correlations were positive and significant, they were not identical, and seldom exceeded .70. This indicates low to moderately strong correlations, indicating that while the questions and categories are related, they are still contributing different and unique information.

Some of the individual questions and categories are not significantly related to each other, whatever their scores might be. This means that a high score on one such question or category gives you no clue as to what the score of the other question or category might be. As expected, none of the index questions or categories were significantly negatively or inversely related, meaning that when the scores of one are low, the others tend to be high.

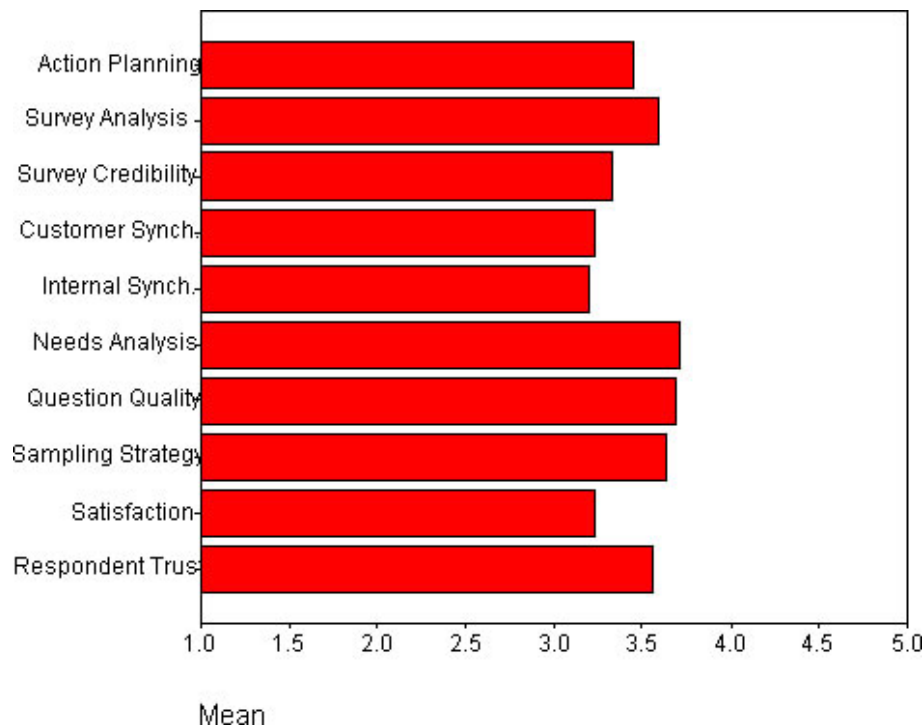
Index Construct

The PDi Survey Effectiveness Assessment Index © contains 35 questions with an agreement rating scale in 10 different categories. The agreement scale featured 5 rating points and a non-response option:

- 1: *Strongly Disagree*
- 2: *Disagree*
- 3: *Neutral*
- 4: *Agree*
- 5: *Strongly Agree*
- 6: *Don't Know*

The 10 categories were established from the 9 factors which emerged from the exploratory factor analysis described above. The satisfaction category questions loaded on the same factor as the action planning category questions, and these categories share the highest correlation. Descriptive statistics and a Pearson Correlation matrix follow the category list.

1. *Needs Analysis*
2. *Sampling Strategy*
3. *Survey Design & Administration: Question Quality*
4. *Survey Design & Administration: Survey Credibility*
5. *Survey Analysis & Feedback*
6. *Action Planning*
7. *Survey Synchronization with Customers*
8. *Survey Synchronization with Internal Partners*
9. *Overall Survey Satisfaction*
10. *Respondent Trust*



Descriptive Statistics

	N	Mean	Std. Deviation
Action Planning	80	3.44	.8057
Survey Analysis	80	3.57	.8498
Survey Credibility	82	3.32	.8204
Customer Synchronization	77	3.21	.9609
Internal Synchronization	79	3.20	.8107
Needs Analysis	81	3.72	.7199
Question Quality	81	3.64	.8148
Sampling Strategy	82	3.64	.7486
Satisfaction	82	3.15	.8841
Respondent Trust	79	3.58	1.1391
Valid N (listwise)	74		

Correlations

		A	B	C	D	E	F	G	H	I	J
Pearson Correlation	A. Action Planning	1.0	.68**	.58**	.39**	.48**	.58**	.39**	.47**	.70**	.26*
	B. Survey Analysis	.68**	1.0	.47**	.36**	.42**	.41**	.41**	.44**	.58**	.27*
	C. Survey Credibility	.58**	.47**	1.0	.44**	.31**	.52**	.42**	.34**	.42**	.43**
	D. Customer Synchronization	.39**	.36**	.44**	1.0	.40**	.32**	.32**	.26*	.38**	.23*
	E. Internal Synchronization	.48**	.42**	.31**	.40**	1.0	.37**	.21	.27*	.43**	.16
	F. Needs Analysis	.58**	.41**	.52**	.32**	.37**	1.0	.36**	.40**	.35**	.20
	G. Question Quality	.39**	.41**	.42**	.32**	.21	.36**	1.0	.33**	.21	.36**
	H. Sampling Strategy	.47**	.44**	.34**	.26*	.27*	.40**	.33**	1.0	.38**	.12
	I. Satisfaction	.70**	.58**	.42**	.38**	.43**	.35**	.21	.38**	1.0	.30**
	J. Respondent Trust	.26*	.27*	.43**	.23*	.16	.20	.36**	.12	.30**	1.0

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Demographic Comparisons

Four demographic categories were included to assess each respondent: gender, ethnicity, managerial responsibility and functional specialization. These demographics were measured by the following nominal questions:

Gender: What is your gender?

- 1. Male
- 2. Female

Ethnicity: Which category best describes your ethnicity?

- 1. Hispanic
- 2. Black
- 3. Asian / Pacific Islander
- 4. White
- 5. Native American / Eskimo
- 6. Other

Managerial Responsibility: Which level best matches your job?

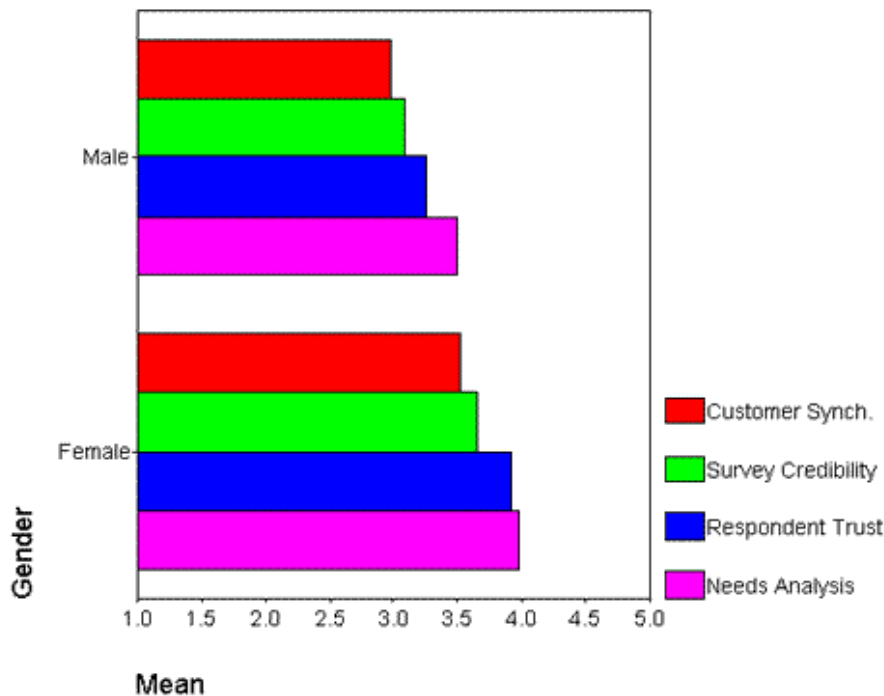
1. No managerial or decision making duties
2. Supervisor - occasionally directs others during a project or task
3. Manager - consistently directs a team or group of people
4. Senior Manager - directs managers
5. Executive - in an executive team of a business unit or autonomous division

Functional Specialization: Which of the following functional areas would you claim as your own, and would claim you as a member?

1. Sales and Marketing
2. Customer Service
3. Product Management (including engineering and development)
4. Operations and Manufacturing
5. Support (MIS, Research, Art & Graphics, etc.)
6. Administration (HR, PR, Legal, Finance & Accounting, etc.)

Most of the demographic groups, with the exception of gender, did not feature enough respondents for reliable comparisons. As the database becomes populated, this problem of small numbers will be resolved.

As for gender, women consistently gave higher ratings on survey effectiveness than men, regardless of the category. These higher ratings were statistically significant (over 99% confidence, under .01 significance) for the survey credibility in the design and administration process, survey synchronization with customers, needs analysis and respondent trust. These differences can be illustrated:



Descriptives

			N	Mean	Std. Deviation
Survey Credibility	Gender	Male	43	3.03	.8349
		Female	39	3.63	.6871
		Total	82	3.32	.8204
Customer Synchronization	Gender	Male	42	2.94	1.0012
		Female	35	3.53	.8130
		Total	77	3.21	.9609
Respondent Trust	Gender	Male	40	3.25	1.2142
		Female	39	3.92	.9565
		Total	79	3.58	1.1391
Needs Analysis	Gender	Male	42	3.48	.7137
		Female	39	3.98	.6387
		Total	81	3.72	.7199

Content Analysis

The PDi Survey Effectiveness Index© pre-test also included three open ended questions. A detailed Content Analysis of the responses to each of these three open ended questions was conducted. The following thematic patterns emerged:

1. What are the primary goals of the survey? Were these goals met?

<i>Primary Goals</i>	<i>Frequency</i>
Improve the Work Place (ID Strengths/Weaknesses)	14
Obtain Employee Input	8
Measure Employee Satisfaction	8
Improve Customer Services	6
Evaluate Management	3
Miscellaneous	7
Total	46
Were these goals met?	
Yes	12
No	6
Somewhat	3
Unsure	1

2. What are the biggest barriers and limitations to effectively using this kind of survey research in your organization?

<i>Barrier</i>	<i>Frequency</i>
----------------	------------------

Poor Survey Format (Design, Questions & Administration)	12
Employees Not Motivated to Participate	8
Confidentiality Concerns	5
No Follow Up/Action Planning	5
Poor Sampling of all Employees	4
Inadequate Resources (Time & Money)	4
No Barriers	4
Results Not Communicated	2
Miscellaneous	8
Total	52

3. Does this organization support survey research well? How could the organization do better?

<i>Theme</i>	<i>Frequency</i>
Improve Follow Up/Action Planning	15
Improve Survey Design, Questions & Administration	3
Improve Response Rates	3
Improve Confidentiality	2
Improve Management Support	2
Miscellaneous	10
Total	35

Content Analysis Summary

The open ended two part question on goals and goal attainment suggests that improving the work place, obtaining employee input and measuring employee satisfaction are the top three goals for conducting surveys in this sample. Given that most of the surveys analyzed were already identified employee surveys, this data did not add enough information to justify the inclusion of this qualitative question in the final survey instrument. Goal realization was mixed, with survey efforts failing to meet their goals 43 percent of the time. Given that an agreement scale would yield more differentiation than the trichotomous themes (yes, no, somewhat) in the written comments, this question was rewritten for use with the 5 point agreement rating scale.

The responses to the open-ended question on barriers and limitations focused on themes around poor survey format, non-participative / unmotivated respondents, confidentiality concerns and lack of follow up/action planning. Similarly, the responses to the open ended question on survey effort support focused on themes around improving follow up/action planning, improving survey questions and improving response rates. Respondents provided minimal data on how their organizations support survey research.

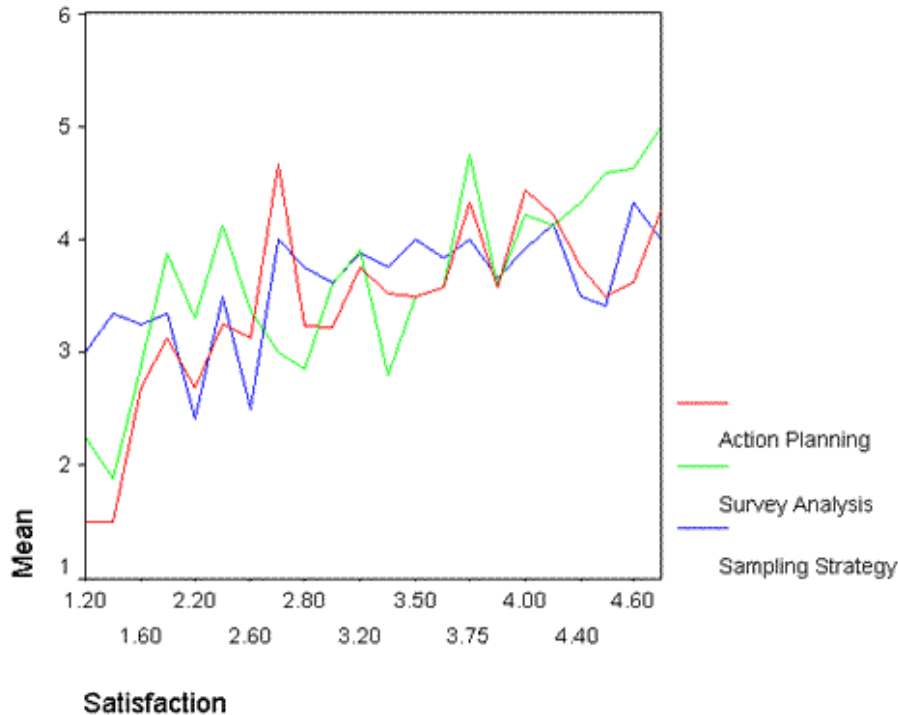
Since respondents focused on the strengths and weaknesses of survey efforts, the open-ended questions were reworded to directly assess these issues. In brief, the initial three open ended questions provided us with the input to better evaluate responses, thus we have reduced the initial three open ended questions to two open ended question that asked about strengths and weaknesses.

Outcome Drivers

The outcome questions assessed overall satisfaction with the quality and value of the survey effort, as well as whether participants would recommend this survey to others. In factor analysis, one of the outcome questions, dealing with participant confidence that guarantees of confidentiality will be honored, loaded on its own factor. Consequently there are two outcome categories: Overall satisfaction and respondent trust.

Survey Satisfaction. Overall satisfaction was positively and significantly correlated to almost all of the other categories, except question quality. This means that when the category ratings were high, ratings of satisfaction also tended to be high. When any of the categories were not rated highly, the

ratings of satisfaction tended to suffer as well. This relationship can be illustrated

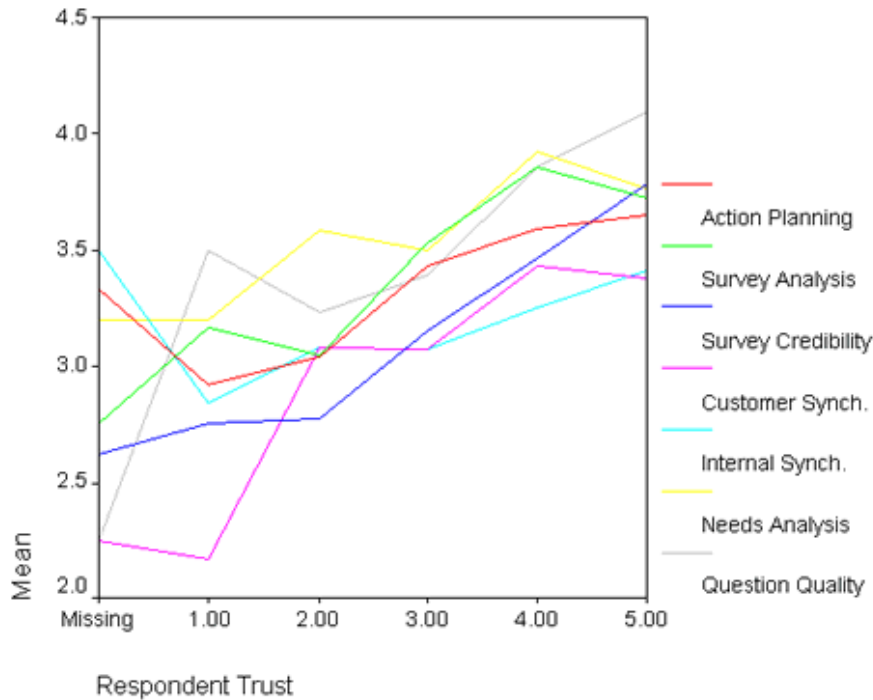


Not all significant correlations have the same explanatory power, however. The higher the correlation, the more ratings of that particular category explain the variation seen in the ratings of a particular outcome question. When a correlation exceeds .50, it is considered a strong correlation. While correlations do not measure causality, change agents who want to improve ratings of a particular outcome variable often target behaviors which are strongly related to that outcome. Chances are that improvements in those areas will improve the outcome of interest as well. Accordingly, categories with the highest correlations with outcomes can be considered drivers.

Of all the categories, action planning is the most highly correlated (.70), to the extent that the action planning and satisfaction questions loaded together on the same factor in factor analysis. This finding is consistent with the high frequency of written comment themes emphasizing the importance of improving follow through and action planning.

Survey analysis and feedback also featured a strong correlation (.52) with satisfaction. Respondents clearly link satisfaction with survey efforts that effectively identify, analyze, and communicate important information and trends. Written comments supported these findings, with respondents targeting poor survey design and reporting issues as a serious weakness. Poor survey quality and credibility issues may explain the written comments noting poor respondent motivation and participation in survey efforts.

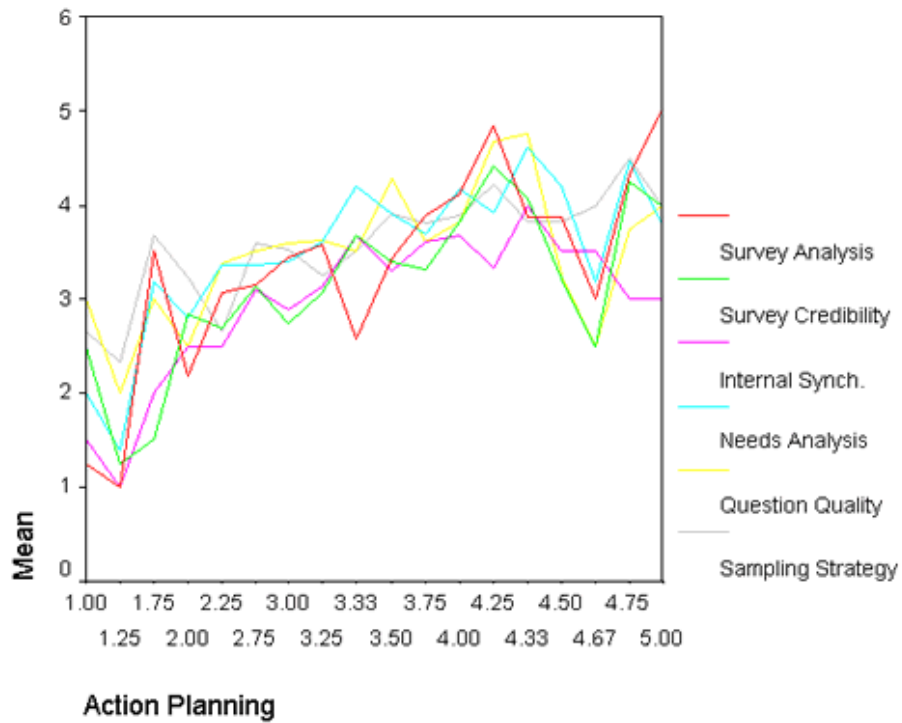
Respondent Trust. This single item indicator assessed whether respondents trusted guarantees of confidentiality. Written comments in response to both the weakness and improvement questions also emphasized the importance of this issue. Trust was positively and significantly correlated to almost all of the other categories, except question sampling strategy and internal synchronization of survey efforts. This means that when the category ratings were high, ratings of respondent trust also tended to be high. When any of the categories were not rated highly, the ratings of respondent trust tended to suffer as well. This relationship can be illustrated:



Of these categories, the strongest correlations with trust were survey credibility (.43) and question quality (.36). This indicates that respondents are most likely to trust promises and guarantees of confidentiality from management when they see quality and integrity in the survey design and administration process.

Implications

Of all the categories, action planning was not only strongly correlated with the satisfaction outcomes (.70), it consistently featured the highest correlations with the other non-outcome categories as well (correlations from .39 to .68). As one of the most frequently mentioned qualitative themes, the importance of this issue in the mind of survey respondents is indisputable. Consequently, in terms of where to prioritize investments to improve the survey process, action planning is an excellent place to start. These correlations between action planning and most of the other categories are pictured below:



[Contact PDi](#) about these publications

(copyright 2002 by PDi. All rights reserved. Please do not modify, copy or distribute this article -- see website [terms and conditions of use.](#)) 