

TeamView/360™ and the Individual Behavior Questionnaire

Psychometric Properties

Kenneth R. Brousseau, Ph.D.

© 1994 Advanced Teamware, Inc. Agoura Hills, CA. All rights reserved. TeamView/360 is a registered trademark of Advanced Teamware, Inc.

TeamView/360 is a software system for processing and presenting individuals' effectiveness ratings as perceived by self and others on a broad range of behaviors. The system also produces team profiles based on aggregated self and peer ratings for groups of co-workers.

The underlying instrument on which TeamView/360 results are based is the *Individual Behavior Questionnaire (IBQ)* developed by Kenneth R. Brousseau and Michael R. Perrault. The IBQ assesses behavioral effectiveness in thirty-one categories, ranging from those that involve primarily cognitive activity to those that involve primarily interpersonal activity. Based on an *a priori* conceptual schema, the thirty-one behaviors are grouped into seven categories. TeamView/360 graphically presents assessment results for the seven categories and separately for the behavioral items in each category.

The IBQ is designed to be a tool to assist individuals in learning about how they are perceived by others with whom they work. The instrument is evaluative, not descriptive. That is, the instrument is designed to assess how effectively the individual is perceived to perform by others and then to enable direct comparisons of others' perceptions of one's effectiveness with one's own self-assessment.

The instrument is not intended to evaluate performance as might be measured by objective outcome measures (e.g., measured number of product defects as an indication of Producing High Quality Performance). This is an important point, inasmuch as there clearly exist situations where peer perceptions could differ from objective reality. For instance, an individual may be perceived as a poor listener when in fact he or she listens well as might be measured by a test of comprehension of others' spoken messages. This kind of situation could arise when others attribute poor listening to another based upon behavioral cues (e.g., not maintaining eye contact when others speak) that often are associated with poor listening. Nonetheless, it generally is important for individuals to

understand what attributions others are assigning to their behavior even when B perhaps, especially when B those attributions are incorrect.

To qualify as a valid measure of others' perceptions, the IBQ must pass two crucial psychometric tests. It must be shown to possess adequate construct validity. That is, it must be shown to measure the particular perceptions that it purports to measure. For example, there should be evidence that when the IBQ assesses perceived effectiveness in Planning for the Future the instrument actually measures how effectively individuals are perceived to be in laying plans for the future and not some unrelated behavior, such as monitoring progress toward meeting established objectives. Secondly, the IBQ must be shown to possess adequate inter-rater reliability. That is, there must be evidence that different raters fundamentally agree in their ratings of the same individual. If the IBQ can be shown to measure the particular perceptions it is intended to measure and that different raters agree in their ratings of the same individual, then the instrument can be considered to be a valid measure of peers' and others' perceptions.

Construct Validity

Factor analysis is a common statistical test employed to assess construct validity. This method evaluates intercorrelations among test items to determine what relationships exist within a set of items. When a factor matrix is rotated by the varimax method, the resulting factors usually indicate how the items in the set group conceptually. Items that show high correlations (i.e., factor loadings) with a particular varimax factor are interpreted as sharing a common characteristic that underlies all items that correlate highly with that factor. A look at the items with high loadings on the factor often indicates the nature of the underlying factor.

A factor analysis of the IBQ should produce a factor solution that corresponds fairly well with the original conceptual grouping of the items into seven categories. Inasmuch as the IBQ contains only thirty-one items, it is unlikely that seven independent factors would emerge. However, the items in each grouping ideally should show their highest loadings on the same factor or, at the least, on conceptually similar factors.

Table 1 shows a varimax factor solution for an IBQ data set, computed by Douglas Saddler and Mildred Murray-Ward of California Lutheran University. The data set contains 4,013 Other Ratings (peer ratings) on each item. This data base represents teammate ratings of about 525 individuals (each individual was rated by an average of 8 or so teammates). About 90 percent of the ratees were middle level to senior level managers; the remainder represents individuals in diverse professional categories -- e.g., nurses, research scientists, investment bankers, technical analysts. Many industries are represented, from aerospace to pharmaceutical to heavy manufacturing to computer products.

The items in Table 1 are grouped into the seven original categories, arranged top to bottom. Each category is bounded by double lines. Proceeding from top to bottom, the

categories are Problem-Solving, Planning, Controlling, Managing Self, Managing Relationships, Leading, and Communicating.

The cells shown in reverse contrast, white text on black background, indicate the highest factor loading for each behavioral item. Cells shaded gray indicate the second highest factor loading for each behavioral item.

Looking at the table we can see that items in the Problem-Solving category all load most highly on the same factor. So do items in the Planning category, the Controlling category, and the Managing Relationships category. However, the data also show that rather than forming two independent factors, the Problem-Solving and Planning items all load on one factor. This indicates that the behaviors in these two categories do not fall into two conceptually distinct categories in the perceptions of raters.

The results for the remaining three categories are mixed. The most problematic category is Managing Self. Two of the items, Handling Pressure and Coping with Own Frustration appear to form their own factor. The item, Self-Development, loads most highly on the Problem-Solving and Planning factor and secondarily on the Controlling factor. The loadings for this item are the least intuitively reasonable among the thirty-one items. Responding to Feedback loads highest on the Managing Relationships factor. Not surprisingly, one of the characteristics of people who are seen as managing their relationships well is that they are viewed as responsive to feedback from others.

The results for the Leadership category show some interesting patterns. Three of the five items load most highly on the same factor. Judging from the loadings of items in other categories that also load on this factor, this is indeed a leadership factor. The list of Leadership items loading on this factor expands to four when secondary loadings are considered

Table 1
Varimax Factor Structure
Rotated Factor Loadings (X 1000)
IBQ Other Ratings (n=4,013)

		Factors				
Item No.	Behavioral Descriptor	1	2	3	4	5
5	Recognizing Trends	-567	236	181	190	221
8	Generating New Ideas	-694	183	140	74	210
30	Evaluating & Acting on New Ideas	-648	285	222	131	182

2	Planning for Future	-551	122	330	138	352
24	Adapting to change	-477	353	154	350	154
3	Organizing Events	-336	211	440	115	344
9	Monitoring Performance	-267	196	470	113	391
11	Meeting Schedules & Deadlines	-125	205	760	176	159
15	Producing High Quality Work	-446	227	555	59	148
19	Maintaining High Productivity	-416	230	563	109	178
25	Meeting Commitments	-204	229	758	181	158
6	Handling Pressure	-237	276	197	656	137
12	Coping with Frustration	-154	321	184	672	139
20	Self Development	-510	227	277	195	216
27	Responding to Feedback	-253	550	225	234	255
7	Initiating Relationships	-264	513	79	183	228
10	Cooperating	-87	710	253	212	46
14	Sizing up People	-281	399	198	184	343
22	Maintaining Relationships	-209	734	127	132	192
26	Resolving Conflicts	-206	482	193	303	313
31	Responding to Others' Needs	-205	691	265	75	181
16	Delegating	-241	140	157	213	505
18	Facilitating Meetings	-361	238	214	104	446
21	Motivating and Inspiring Others	-398	546	169	123	367
23	Developing People	-338	447	138	105	515
28	Giving Recognition	-261	535	165	52	356
1	Articulating Ideas	-471	217	251	120	316
4	Listening	-170	637	155	218	140
13	Keeping Others Informed	-166	447	391	100	306

17	Giving Feedback	-205	319	215	69	619
29	Communicating Expectations	-335	319	260	63	49

Interestingly, three of the five items have high secondary loadings on the Problem-Solving/Planning factor.

This indicates that although leadership behaviors do tend to fall into their own category, they share a clear, intellectual component. These results seem to indicate that to be viewed as a competent leader requires intellectual skills at least as much as, if not more than, interpersonal skills.

Looking at the Communicating category, we see a clear pattern. Even though the highest loadings in this category vary from one factor to another, four out of five of the items in this category have their highest secondary loadings on the same factor as do those in the Managing Relationships category. The indication is that Communication behaviors differ from one another considerably, but nonetheless do share similarity in the Relationships category. Standing back further, looking at all of the loadings across categories for Factor 2, on which all of the Managing Relationships factors load most highly, we can see that this factor could be more aptly labeled, General Interpersonal Competence.

Looking at the results in Table 1, we can see that, within categories, an average of 77 percent of the items load most highly on the same factor. When secondary loadings are considered, we see that, within categories, 90 percent of the items load highly on the same factor. Significantly, these results are essentially identical to a prior varimax factor solution computed several years ago with a data set consisting of several hundred, rather than several thousand, cases. These results provide substantial evidence supporting the construct validity of the IBQ.

Inter-Rater Reliability

Table 2 shows the results of inter-rater reliability computations produced by Murray-Ward and Saddler. Working with the same data-base as described above, they partitioned the sample according to the number of raters. This produced 14 partitions (from 2 to 15 raters). Within each partition, they used Ebell's Reliability by Intraclass Correlation formula to calculate inter-rater reliability for each behavioral item. Inspection of the results indicated that the best reliabilities are achieved with between 6 and 11 raters.

The mean inter-rater reliabilities for groups of raters in this size range are shown in Table 2.

Reliabilities tend to fall off with fewer or greater numbers of raters, particularly with fewer than four raters. The lower reliabilities for small groups of raters are not surprising and probably reflect the fact that a single outlier can skew ratings easily.

The lower reliabilities for larger groups of raters is less expected intuitively. However, with larger numbers of raters the probability is raised that the group of raters includes individuals who are not closely familiar with the person being rated, and this likely accounts for the declining reliability coefficients with the larger rater groups.

Table 2 contains ample indications of fully adequate inter-rater reliability of the IBQ overall, indicating that different raters basically agree in their ratings of the same individual on the IBQ behaviors with a mean reliability of .72.

Nonetheless, a few words of caution are in order regarding several items that show considerable lower inter-rater reliability, such as Responding to Feedback, Sizing up People, and Monitoring and Controlling Performance. The data in Table 2 suggest that different raters tend not to agree so strongly in their assessments of individuals' effectiveness in these categories. This may indicate that people have different ideas about what constitutes effective behaviors in these particular categories.

Table2		
TeamView/360 - IBQ Inter-Rater Reliability		
(6 - 11 raters)		
Category	Behavioral Item Descriptor	Mean Reliability Coefficient
Problem-Solving	Recognizing Trends	.68
	Generating New Ideas	.78
	Evaluating & Acting on New Ideas	.72
Planning	Planning for Future	.80
	Adapting to change	.75
Controlling Work	Organizing Events	.76
	Monitoring and Controlling Performance	.60
	Meeting Schedules & Deadlines	.78
	Producing High Quality Work	.74
	Maintaining High Productivity	.73
	Meeting Commitments	.76

Managing Self	Handling Pressure	.71
	Coping with Frustration	.69
	Developing Own Capabilities	.72
	Responding to Feedback	.59
Managing Relationships	Initiating Relationships	.73
	Cooperating	.77
	Sizing up People	.59
	Maintaining Relationships	.75
	Resolving Conflicts	.62
	Responding to Others' Needs	.74
Leading	Delegating responsibility	.74
	Facilitating Meetings	.75
	Motivating and Inspiring Others	.80
	Developing People	.72
	Giving Recognition	.74
Communicating	Articulating Ideas	.79
	Listening to others	.74
	Keeping Others Informed	.68
	Giving Performance Feedback	.64
	Communicating Expectations	.62

Or, it may indicate that people simply find it difficult to estimate effectiveness on these particular dimensions based on observable behavior. Consequently, practitioners should interpret ratings in these categories with caution and only with supporting evidence from other ratings.

Summary

The results of analyses presented in this paper provide ample evidence supporting the use of the IBQ and the associated TeamView/360 system for purposes of presenting individuals with valid indications of how they are perceived by work associates.

The IBQ appears to have a conceptual sound factor structure that agrees substantially with the original conceptual grouping of the behavioral items proposed by the authors. Moreover, the results of inter-rater reliability analyses based on several thousand sets of IBQ ratings indicates acceptable levels of agreement among raters in how they rate the same individual.

Nonetheless, the results do suggest that IBQ ratings should be interpreted with care in certain categories. First of all, overall summary scores of effectiveness in the Self-Management category should not be relied upon in view of the fact that items within this category do not cluster well. Instead, results of ratings of each behavior should be the focus of attention in this category. The inter-rater reliability results for individual behavioral items indicate good reliability with the one exception of Responding to Feedback, so the results warrant use of items in this category without reference to any overall summary or factor score.

To a much lesser extent, similar caution should be observed in the Leadership category and the Communicating category.