Quantifying Qualitative OD Results: Dispelling the touchy-feely stigma

A Research Paper submitted by

LISA NIELSEN

To

SONOMA STATE UNIVERSITY

In partial fulfillment of the requirements for the degree of

MASTERS OF PSYCHOLOGY

In

ORGANIZATION DEVELOPMENT

This culminating paper has been accepted by the faculty of the Masters Program in Organization Development

Sara Tickler, First Reader
Bo Laurent, Second Reader

Rohnert Park, 2011
Abstract

Evaluation enables agility and ability to self-correct mid-course. It also facilitates OD practitioners with an opportunity to continually learn from their experiences to catalyzing the evolution and refinement of their tools and skills. Oddly, the field of OD does not have an extensive history of evaluation; the quality, scarcity and validity of its evaluations have been heavily questioned since the field’s inception. In contrast to the reflective rigor one might expect from a discipline that advocates high degrees of reflection from its clients, OD evaluations have provided largely anecdotal information; criteria for success has been subjective and testimonial in nature. This phenomenological research paper examines what forms of evaluation current OD practitioners utilize, what is missing from their approach, the implications this has for the longevity of the field of OD, and recommendations for further development in evaluation methodology.
Quantifying OD Results

Introduction

The field of Organization Development (OD) was founded upon the aspirations of innovation and improvement. Curiosity of early pioneers such as Frederick Taylor (1856-1915), Kurt Lewin (1890-1947), Douglass McGregor (1906-1964), Fred Emery (1925-1997), and Eric Trist (1909-1993) drove exploration of methodologies to support organizations to be more effective (Weisbord, 2004). Keeping in mind the spirit of continual improvement, one would anticipate the field of Organization Development to have a highly established and rigorous method to evaluate the effectiveness of its technologies and interventions. As Lewin (1946) himself wrote:

If we cannot judge whether an action has led forward or backward, if we have no criteria for evaluating the relation between effort and achievement, there is nothing to prevent us from making the wrong conclusion and to encourage the wrong work habits. Realistic fact-finding and evaluation is a prerequisite for any learning. (p. 35)

Evaluation provides an opportunity for OD practitioners to continually learn from their experiences and to catalyze the evolution and refinement of their tools and skills. Oddly, the field of OD does not have an extensive history of evaluation; the quality, scarcity and validity of its evaluations have been criticized since the field’s inception by authors such as Bennis (1965), Dunnette (1974), Lippitt, (1969), Morrison (1978), and Pearse (1970). In contrast to the reflective rigor one might expect from a discipline that asks such reflection from its clients, OD evaluations have provided largely anecdotal information; criteria for success has been subjective and testimonial in nature (Hartke & Vicars, 1984).

Relevancy of this Research

The abilities to understand cause and effect (single loop-learning) and to reflect upon and evaluate the underlying assumptions guiding those causational relationships (double-loop learning), are necessary if practitioners and organizations are to make informed decisions in rapidly changing and often uncertain contexts (Argyris & Schon, 1974; 1978; Argyris, 1982; 1990). Just as OD clients are advised to learn from their processes and move with intention rather than reaction, as a field, OD might follow its own sage advice.
Quantifying OD Results

The financial crisis that began in 2007 triggered a global recession in scale and magnitude not seen in nearly a century (World Economic Forum, 2009). Surviving this economy has required of organizations unprecedented levels of innovation and financial prudence. “Global competition, relentless cost cutting and job elimination have elevated the needs of the organization for competitive success over the needs of individuals” (Marshak, 2005, p. 25). Questions of Return on Investment (ROI) are expected dialogues when contemplating business development strategies. Unfortunately, the field of OD is not well prepared to have conversations in this language of business: financial figures, balance sheets, metrics, forecasting and bottom-lines. “OD practitioners do not necessarily need to be experts in business strategy, but they often do not sufficiently understand the language and how profit is made and costs contained according to various business models” (Bradford & Burke, 2005, p. 8).

As a new OD practitioner, I personally find this discouraging because it is my belief that OD interventions have the potential to add great value to organizations’ scorecards. However, without the language and methodologies to quantify OD services, the value is not represented and the credibility of being a revenue-creating, cost-cutting resource is not generated.

There is a common perception of OD as being a soft skill, i.e. communication, conflict resolution, personal mastery, or teamwork tool (Business Dictionary, 2010). And while many organizations regard soft skills as pleasant to have, they are not regarded as critical to the bottom line. I and many other humanistic-oriented practitioners would argue that soft skills are critical to the bottom line because productivity is bound to remain inefficient until the whole system is engaged and unless the synergy of a learning organization is harnessed (Senge, 1990). Weisbord (2004) argues that most people want dignity, meaning and community from work. Workplace environments rich in these qualities can help reduce costly conflict, absenteeism, and turnover. However, despite persuasive arguments for the usefulness of OD’s ability to influence workplace environments, at present, “Executives ignore OD or relegate it to the bowels of the organization” (Bradford & Burke, 2005, p. 7). To reach the executive table where budgets are drawn and strategies planned, OD practitioners must learn to translate their results into the language
of executive decision makers: the language of profit and loss. Otherwise the tremendous value of OD work remains unseen, unheard, and undervalued.

If OD practitioners can’t measure the effectiveness of their work, how can clients put a value on it?

**Personal Relevancy**

I joined the field of OD because of a desire to increase well-being in the lives of individuals. I have previous undergraduate and graduate psychology degrees and for some time contemplated pursuing a private clinical practice. However, I thought about how much time the average person spends at work, often half their waking hours. Indeed, many individuals report their workplace to be a second home. I thought to myself, “why not strive to increase individual well-being in a large-scale systemic way?” Organizations and work-life experience offer tremendous leverage points in increasing the life-satisfaction of many individuals. This realization inspired me to join the field of OD which has valuable technologies to bring to the workplace to increase feelings of autonomy, status, certainty, relatedness, and fairness—five of the most powerful contributors to an overall sense of well-being (Rock, 2009). Not only do these improvements benefit the happiness and physiological health of the employees, but productivity, innovation, accountability, and loyalty are also increased throughout the organization. So why isn’t OD being utilized in every organization, all the time? This query is what led me to the topic of OD’s credibility as a revenue source, and how to establish credibility through the language that is indisputable in the business world: quantitative hard data.

**Purpose of This Research**

The purpose of this research is to discover what types of OD intervention evaluations are currently being utilized within the field of OD and what avenues for further development of evaluation efforts may be available and useful. The research also briefly explores why evaluation has not evolved to be standard procedure in OD and what the arguments against evaluation might be. This research addresses the OD field’s ability to sustain itself in a changing business marketplace.
Literature Review

Despite Lewin’s urging for evaluation in 1946, published literature indicates that early attempts to conduct OD intervention evaluations were not highly successful and received a considerable amount of criticism. According to Vicars and Hartke (1984):

When systematic evaluations have been attempted, for the most part they have had serious methodological weakness: Limited use of control groups, focus on short-term changes, and lack of independent evaluations are three in particular that have been complained about frequently by OD critics (Back, 1974; Dunette, 1974; Morrison, 1978; Pearse, 1970). (p. 177)

Vicars and Hartke’s (1984) research was inspired by an earlier assertion by Morrison (1978) that “The studies of OD evaluations published to date do not conform to established criteria for internal and external validity for effective social science research and evaluation” (p. 65). Vicars and Hartke re-tested current published OD evaluations, using Morrison’s criteria, which were based on Campbell and Stanley’s (1963) Twelve Threats to Validity, as criteria of the quality of evaluation research designs. The 12 threats are:

1. History-Effects of simultaneous events. Control group combats this threat.
2. Maturation- Can the results be attributed to natural development of the group?
3. Testing- Did pre-test and post-test sensitize respondents?
4. Instrumentation-Changes in the measuring instrument?
5. Statistical Regression- Were the study group selected based on extreme scores?
6. Selection Bias- Difference in experimental group and control group?
7. Experimental Mortality- Was there a difference in the participants between the pre and post tests?
8. Selection-Maturation Interaction- Is there any characteristic in the experiment group that is not present in the comparison group that would naturally lead to the changes observed?
9. Interaction Effect of Testing- Did the pre-test make the experiment group more sensitive to the experimental variable?
10. Interaction Effects of Selection Biases and the Experimental Variable- Were there characteristics of the experimental group population that would cause the results of the experiment to differ from those to be expected in the general population?
11. Reactive Effects of Experimental Arrangements- Were there qualities about the setting that made the results not generalizable to a non-experimental setting?
12. Multiple Treatment Interface- Were there other treatments applied to these groups that may have contributed to the effects noted?
The findings of Vicars and Hartke (1984) indicated that while there was a slight improvement in evaluation research design between 1978 and 1984, “Too many of these evaluations have provided only anecdotal information, and criteria for success have been too subjective and testimonial in nature” (p. 177).

An attempt was made to discover, from the perspective of the OD practitioner, what was blocking successful OD evaluations. In *Guidelines for overcoming empirically identified evaluation problems of organizational development change agents*, Armenakis, Field, and Holley (1976) argue that, “The evaluation phase of the OD process is crucial to the OD effort since data from the evaluation serve as inputs to the participants as well as a basis to justify the time and money expanded in the effort” (p. 1149). These authors sent a survey to the 269 current members of the Organization Development Network to ascertain the evaluation challenges of OD practitioners. A total of 107 evaluative challenges were identified and subsequently clustered into three problem categories: methodological, administrative, and miscellaneous. They discovered, “The problem most frequently encountered by these change agents was the difficulty in selecting and quantitatively measuring ‘soft’ criteria” (p. 1151). The second most frequently encountered problem in performing evaluations of OD efforts was the difficulty in employing comparison groups to evaluate change resulting from the OD intervention.

**The Influence of Positivism on Evaluation Methodology**

It may be useful at this juncture to examine the source of evaluation methodology applied to the OD field. As previously stated, Kurt Lewin is attributed with being one of the five most influential pioneers of OD (Weisbord, 2004). “The intellectual foundation for OD began in the 1940’s with the research and writing of Kurt Lewin and his protégées” (Greiner & Cummings, 2005, p. 89). Lewin is also credited as being the founder of Modern Social Psychology (NTL Institute, 2008).

In 1916 Lewin obtained his Ph.D. at the University of Berlin, completing his dissertation under Dr. Carl Stumpf who was interested in empirical research methods in experimental psychology, and who
became one of the pioneers in the empirical research discipline (Stumpf, 1930). In 1944, Lewin was appointed Director of MIT’s newly created Research Center for Group Dynamics (Marrow, 1969).

“Lewin’s life, like Taylor’s, was marked by a passion for experimentation and a faith in science” (Weisbord, 2004, p. 76). This brief history is relevant because OD evaluations, as analyzed and critiqued by the previously mentioned authors, were based on the model of empirical testing, as developed by the psychologists (a discipline that grew out of the medical field), such as Carl Stumpf in the 20th century.

Empirical testing can trace its roots back over a thousand years to the scientific method as originated in the time of Aristotle, further developed by Ibn al-Haytham and solidified by Galileo Galilei (Moulton & Schifferes, 1960; Kline, 1985). Much of the orientation of the scientific method can be summarized by the philosophy of positivism, a belief that only that which is based on sense experience and positive verification can be considered authentic knowledge.

Among the many streams contributing to the original formulation of OD was a strong positivist orientation rooted in mid-twentieth-century social science research methodologies. The whole premise of data-based change (for example, action research and survey research methods) presumes the existence and validity of an objective, discernable reality. (Marshak, 2005, p. 25)

Although there are no fewer than twelve distinct epistemologies that are referred to as positivism, several principles run throughout:

- The unity of the scientific method— the logic of inquiry is the same across all sciences (social and natural) i.e. a single experiment design should work across disciplines, be it medicine, biology, psychology, or organization development.

- The goal of inquiry is to explain and predict. Most positivists would also contend that the ultimate goal is to develop the law of general understanding, by discovering necessary and sufficient conditions for any phenomenon (creating a perfect model of it). If the law is known, we can manipulate the conditions to produce the predicted result- i.e. if conditions are the same, an OD intervention should produce the same result each time it is applied.

- Scientific knowledge is testable. Research can be proved only by empirical means, not argumentations. Research should be observable with the human senses (arguments are not enough, sheer belief is out of the question). Positivists should prove their research using the logic of confirmation- i.e. personal accounts such as “I feel more satisfied at work” are not scientific evidence. Results must be observable, quantifiable, and able to be verified via third party. (Halfpenny, 1982).
The essential summary of the above principles is that reality and truth exist, in absolute terms, and if a scientist or researcher can isolate contributing factors, cause and effect can be determined and accurate predictions made, i.e., if you take blood pressure medication, your blood pressure will go down. In OD’s case the positivist scientific method would enable evaluations to determine that if you conduct a process-mapping OD intervention, productivity will go up by x%.

**Research Focus**

Since much of modern OD practice is based on post-modern thought and theories of social constructionism which stand in stark contrast to positivism, this research paper asks: Is there a methodology to evaluate OD interventions that would suit the discipline more aptly than traditional empirical testing?

The added benefit to an evaluative technique that fits OD might be, as Beer (1976) writes, “Developing a method for measuring these effects can not only increase the effectiveness and professional responsibility of OD but it can increase the power of OD within organizations” (p. 50).
Methodology

Research Design

A qualitative, phenomenological research method was used to investigate what type of evaluations current OD practitioners utilize to measure the impact of their interventions within organizations. Research participants were pursued through a purposeful selection method, based on research criteria, to form an eight person panel. The interviews conducted provide the primary data for analysis and interpretation regarding the topic of evaluation methods. Research criteria for panel selection included practitioners who:

- Have practiced a form of Organization Development (OD), Organization Effectiveness (OE), or Change Management (CM) for a minimum of five years.
- Work as an internal consultant in a for-profit organization, or as an external consultant.
- Ideally demonstrate some kind of financial impact or value-add for their OD work.

Participants

The panel consisted of eight participants with experience in the field ranging from five to 30 years. The average tenure was 15 years. Fifty percent of the respondents have worked both as external and internal consultants, while 50% have worked solely as internals. Three women and five men were interviewed. All of the internal respondents are employed by for-profit organizations with national and/or global offices including Kaiser Permanente, Motorola, Agilent (Previously HP), Fosters Estates, Implementation Management Associates, and Boeing. One of the participants works as an outward facing consultant for one of the “big four” multi-national accounting and advisory firms. While the participant pool is not considered a representative sample, their combined experience offers a rich depth of knowledge and breadth of experience of the Organization Development field.

Process

Participants were identified through direct personal contact, or indirectly through colleague referrals. Four of the interviews were conducted in person and four via telephone. In-person interviews were tape-recorded, and key quotes were later transcribed. Phone interviews were not recorded, though
detailed electronic notes were taken. All of the participants were asked the same 15 interview questions, however differing follow-up questions were asked of participants for clarification purposes. All interviews lasted approximately one hour. For a list of interview questions see Appendix A.

**Data Analysis**

Analysis of the data was pursued through a coding and thematic analysis approach. The design strategy for developing substantive thematic category topics during the data collection process was inspired by the Action Research Model. After completing the interviews, ten quotes were selected from each interview that captured the speaker’s salient points and were able to stand alone and hold meaning. These quotes were placed on an anonymous quote wall, color-coded and grouped into related categories so that as the data accumulated, the researcher was able to identify emerging cross-subject themes for later analysis and discussion. For an image of the coding methodology see Appendix B.

**Research Augmentation**

In addition to the phenomenological interview data collection, a ten question anonymous online survey was designed and distributed to alumni and faculty of the Sonoma State Organization Development Masters Degree program as well as the Fabulous Facilitators professional network, which is described as a “learning community of seasoned facilitators and those new to the field,” and was founded in 2002. All recipients were also requested to forward the survey on to any additional professional networks. Fifty-four voluntary respondents replied. For a list of the survey questions see Appendix C.

**Assumptions**

Assumptions made in pursuit of this research and while analyzing the data included the following:

- OD interventions have the potential to increase organizational revenue by increasing innovation and efficiency from staff and/or to reduce costs such as attrition, failed mergers and failed change projects.
- Some OD interventions are not successful, or not as successful as they could be, and these occurrences are not being critically documented to be learned from and improved
Quantifying OD Results

upon. Although this phenomenon is not exclusive to the OD field, it hampers the improvement trajectory.

Biases/Limitations

An intentional limit on the participant interview pool was created with a boundary drawn around internal for-profit practitioners. While the external practitioners selected may have consulted to non-profit organizations at some point in their career, the objective was to pursue mostly for-profit organization engagements. The assumption is that due to their business model orientation, for-profit organizations are inherently more concerned with profit margins and return on investment and thus more prone to desire an analysis of the effectiveness of the OD intervention.

A limitation of the anonymous survey is that results were mostly sourced from a Sonoma State University alumni pool and non-university affiliated California Bay Area practitioners. While respondents were encouraged to forward the survey on to professional colleagues to increase diversity of the respondent pool, this common training and/or geographical proximity could influence practitioner habits and thus survey results.

A third limitation important to acknowledge is the possible variation of the interviewees’ clients’ pre-existing perceptions of the credibility of Organization Development. Pre-existing client perceptions may affect the OD practitioners’ perceived credibility and greatly influence the need or desire to perform (or not) post-engagement evaluations.

Lastly, it is important to note the researcher’s bias to see the field of OD remain relevant and respected within the business community. This research was theoretically framed by the questions, “How can humanistic-oriented disciplines such as OD assume their place in the strategic executive team-- not simply to facilitate the strategic planning sessions as an objective process guide, but as an employed participant. How can humanistic practices be demonstrated to be a revenue-source, rather than cost center?” This bias may influence the data gathering and analysis through unintentional preference.
Results

As mentioned previously, each participant was asked the same set of interview questions.

The interview questions focused on four areas:

1) The respondent’s background and the level at which they typically enter organizations
2) The discussion of ‘success expectations’ with new clients during the contracting phase
3) Post-engagement evaluation habits
4) Reflective questions regarding establishing credibility with new clients

Some of the interview participants were clear supporters of metrics and evaluation practices, while some did not seem to feel they were necessary, especially in regard to establishing credibility with their clients. Initially, interviews were pursued to discover what types of evaluations were being performed by successful OD consultants, specifically chosen for their involvement in large, metric-centric corporations, influenced by a hypothesis that these practitioners may be the pioneers who could point toward evaluation technologies. However, what was discovered during the interviews is that evaluation practices, even amongst this esteemed coterie, are scarce or subjective in nature. The major themes that arose from the interviews are:

- What evaluation methods are being utilized?
  - Existing metrics
  - Tracking milestones
- What gets in the way?
  - Evaluation is complicated, time consuming and expensive.
  - It’s the client’s job.
- What is done instead?
Quantifying OD Results

- Demonstrate business acumen and language.
- Form “partnerships” with clients.

Existing Metrics

Often times, the first thing respondents identified when speaking of methods they used to evaluate the results of their interventions was reliance upon pre-existing metrics within the organization such as sales figures, employee satisfaction survey results, inventory numbers, employee retention or attrition rates, leadership scores, cycle-times, and so forth. As one internal interviewee states:

 Typical, projects come up because there is some sort of data telling us there is a problem. We have quality metrics, financial, customer satisfaction, employee engagement, speed and decisiveness, customer focus (connection with customer), dashboards and quadrants, surveys; we measure all of that on a monthly basis, we report that out, we are very transparent, monthly and quarterly. When we can start to see the indications that something is going off the rails, that’s when we choose to intervene. By starting with metrics on a regular basis, we can manage our business effectively, so we know how to modulate what we are doing well (Respondent#1). From here forth, ‘Respondent’ will be shortened to “R.”

While not all the respondents interact with organizations that track figures so rigorously, most shared the sentiment that “ideally the client has a set of metrics in place we can use” (R#7). Another interviewee discussed this benefit: “Projects that can leverage existing metrics within the organization, such as inventory on hand/unused/cost of keeping, are useful; having a hard measure gets the client intrigued” (R#3). Part of the reason existing metrics were described to be ideal is because they are already being collected by the client, so it eliminates designing and delivering an extensive evaluation process, and secondly because if the client is already tracking these figures, the consultants presume these figures are inherently important to the client and/or their organization, and thus the client will be more likely to be engaged.
Tracking Milestones

In addition to, or in place of, utilizing existing metrics, interview respondents stated that they often rely upon achievement of key milestones or deliverables to track the success of a project. Sometimes “evaluation” is simply a question of, “Did it get done?” “Was it adopted?” or “Was a new process implemented?” Commonly indicated milestones include “typical life-cycle stages: design phase, delivery, implementation, modification” (R#7) and examining whether these stages were reached and completed on a project. One interviewee explained:

For a lot of projects clients want relief, improvement…They feel pain or a need, but they haven’t yet articulated what they want as an outcome, they just know they want change and out of the pain…. They want to have a plan as a result of your work with them. They want to be able to make decisions more easily, or establish a process that everyone understands and follows (R#5).

Another consultant states that she tracks “dates, processes, core deliverables such as the creation of a revenue model, or completion of an environmental scan. Success is gauged on how it moved their thinking-- that’s not quantifiable. Also, was it completed on time and accepted by the client?” (R#4) The reason practitioners track if it was done, not how well was done is explored in the next section theme.

Evaluation is Complicated, Time Consuming and Expensive

Frequently, when asked about evaluation techniques, interviewees instead responded with the reasons evaluation does not always happen. All of the interviewees shared the observation of R#2, a long-term consultant, that “One of the problems with measurement is that there are so many other contributing factors.” R#1 expanded on this idea, “What I do is all in service of the business, so it’s not independent of what the business is doing. I am embedded, so I am not measuring the ROI of each project I am doing. My work is more integrated into the overall ROI of the business.” This concept is easily summarized by the perspective, “If you are doing a project within one department, all the business still affects the end product” (R#3).

Another frequently cited reason for a lack of standardized evaluation practices is that “we try and use what is pre-existing, but sometimes it is unusable. Sometimes we have to seek out our own data, but
that process is time consuming and expensive‖ (R#7). Agreeing with the challenges of time and money restrains, another consultant embellished on the challenges, ‘People are just not going to spend a lot of money on measurement because senior leaders are all about installation and then on to the next thing. They get reinforcement for installation. Strategy design is much sexier than measurement‖(R#6).

Apparently, ―taking the time to go back and evaluate is hard, seems like businesses just want to move on to the next thing‖ (R#3).

It’s the Client’s Job

An additional theory was offered by several interviewees that evaluation of the effectiveness of OD interventions is “nearly impossible because I’ve stepped into what they [the client] want to take credit for‖ (R#4).

I’ve had clients who don’t want to go to the ROI conversation, saying, ‘that’s our job.’ Usually my work enables them to get that performance they want to measure. I might be building the sales management team, or plan and do performance management more effectively, so I might be one degree of separation from that hard measure they are after (R#5).

Not all of the interviewees indicated that they experienced evaluation resistance from the client due to the client’s wanting to take credit for the results. Many respondents, however—particularly the external consultants—expressed that the evaluation is the responsibility of the client; “they may handle that part internally” (R#7). Or put more strongly, “that would be more for the client to do. If they don’t identify desired behavioral changes up front, they can’t measure adoption at the back end. And since most don’t, they are not in a position to measure. They [the client] measure success based on installation” (R#6). One consultant stated that every large-scale change engagement project ended with a standardized evaluation: a survey completed by the client. “There is a ‘project value’ assigned by the client for what they think this project is worth to their bottom line. I’d like to think this number is arrived at via hard data, but it is a client generated number and I tend to think it is rather subjective (R#8).”

Demonstrate Business Acumen and Language
Since evaluation was largely described to be inconsistent or problematic, what happens instead? How do these consultants establish credibility with their clients? Many of the interviewees described their ability to connect with their clients and establish rapport through their initial scoping conversations where they believed they demonstrated to their potential clients that they understood their business concern or impediment and were able to knowledgably converse about solutions. The general sentiment of the interviewees was, “I’ve got to be able to speak their [clients’] language and understand what their issues are. I don’t need to be a technical expert, but I need familiarity” (R#8). One candidate described the leverage business acumen provides during initial contracting:

I can speak their language. I have to know what I am talking about, specific to their business. I have to have enough of a background in business to have conversations when I am at the table. OD can be imbedded in so many parts of the business, but if I am looking at a P&L [Profit and Loss] I can say, ‘it looks like you have a problem here, maybe I can help you with that’ (R#1).

Another practitioner vocalized a concern about a lack of business language within the field and the advantage a business approach provides in accessing top decision makers. “I think too many people in the field are not business people and that causes a huge amount of problems for the field; they don’t have the language. We are really about being business people...we don’t want to be seen as soft. I think it helps get the leader’s attention” (R#6). One interviewee described the type of language he might use with senior leaders: “They know hard financial measures, balanced score card, continuous improvement measures. I like to use this as a basis: earnings before income tax or expense ratios. What they don’t often see is the human side: what is it really going to mean if we align customers in decision making around product design and enhancement? Those are the things I try and help them find key indicators around” (R#2).

Along the lines of building validity, one consultant stated, “Being able to talk through the details of how you designed and implemented past projects builds trust and validity” (R#3).

**Form Partnerships with Clients**
Part of what was often described by respondents during the interviews was the propensity to continually check-in with the client throughout the engagement asking “is this on target?” Since many of the practitioners stated that they are tracking success by whether or not a milestone was delivered, not necessarily trying to measure its effectiveness, they rely upon the client’s subjective sense of the project’s usefulness. “As we build the process, it is checking back in with the client, ‘Is this working?’ ‘Is this going to get you where you want to be?’ ‘Is this what you want to be doing?’ A lot of it is constant renegotiation and communication because the objective changes” (R#5). Almost every interviewee echoed this sentiment: “If I start with, ‘how can I help you be successful?’ people want to work with that” (R#1).

Research Augmentation Results

Much of what was discovered during the interviews also surfaced in the anonymous survey results. Thirty-nine percent of survey respondents self-identified as OD practitioners, 23% as Business Consultants, while 27% stated the second most common professional title they identify with is Change Management; 24% stated their second most commonly used professional title is Training and Development. Figure 1 demonstrates the tenure of survey respondents.

<table>
<thead>
<tr>
<th>Figure 1: How long have you been working as a practitioner?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Years</strong></td>
</tr>
<tr>
<td>0-3 years</td>
</tr>
<tr>
<td>4.5 years</td>
</tr>
<tr>
<td>6-10 years</td>
</tr>
<tr>
<td>11-15 years</td>
</tr>
<tr>
<td>16-20 years</td>
</tr>
<tr>
<td>21-25 years</td>
</tr>
<tr>
<td>26-30 years</td>
</tr>
<tr>
<td>31+ years</td>
</tr>
</tbody>
</table>
Seventeen percent of the respondents reported that the majority of their work has been as an internal consultant. Fifty-nine percent report that the majority of their consulting work has been external and 24% stated that they have consulted in a combination of internal and external roles.

When asked how frequently they discuss expectations for success with their clients during the contracting phase of their engagements, 63% reported that they always discuss expectations for success; however only 13% responded that these discussions include hard data. The largest majority, 33% reported that hard data is rarely discussed in terms of success expectations (see Figure 2).

When asked what method practitioners use to evaluate the results of their projects the most frequently used form of analysis was anecdotal conversations, with 90% of practitioners reporting they...
utilized this method. The second most commonly used method of post-engagement analysis was via survey and/or questionnaires, with 65% of practitioners reporting they utilize this method. Thirty-one percent of respondents report relying upon common business metrics to analyze results and 48% state that they create their own metrics. Zero percent of practitioners reported utilizing Campell and Stanley’s (1965) Twelve Threats to Validity Guidelines (see Figure 3).

When asked how practitioners establish credibility with new clients, 35% of respondents ranked the most important aspect of gaining credibility as having a referral. Thirty-three percent reported that credibility is best established by asking the right questions, while only 11% believed having financial figures or metrics of previous engagements is essential to establishing credibility with new clients (see Figure 4).
Figure 4: How do you establish credibility when contracting with a new client?

- By forecasting probability of success: 4%
- By demonstrating business acumen: 13%
- Word of mouth / referrals: 35%
- By asking the right questions: 33%
- Discussing previous anecdotal success: 11%
- With financial figures or metrics of previous engagements: 4%
Discussion

Learning Implications of Current Practitioner Evaluation Methodology

As evidenced in the data, the majority of practitioner respondents rely upon anecdotal conversations and subjective surveys as evaluative measures, which is strikingly similar to the research findings of Vicars and Hartke (1984). While this data pool is limited and not considered to be representative of the entire population of OD practitioners, the data is consistent with current criticisms and concerns of OD practices, many of which were cited in the Literature Review. The apparent lack of objective analysis translated into the business language of metrics and hard data reveals how OD acquired the stigma of being “soft.” As Robert Marshak (2005) writes, “Given the language, values, and ideological orientation of OD, it should not be surprising that it has been labeled by its critics as ‘too touchy-feely’” (p. 25).

Not evaluating results also inhibits an opportunity to engage in double-loop learning about the effectiveness of OD interventions. While OD practitioners often report forming thought-partnerships with their clients to inquire, “Is this it; are we getting it?” which is a form of reflection, it exists within a dangerously limited realm of subjectivity and personal experience. Chris Argyris and Donald Schon (1974) pioneered the concept of double-loop learning. Argyris (2005) explains the risk of subjective reflection, defining it as defensive reasoning. He writes (2005):

Defensive reasoning cannot be objectively verified… A defensive reasoning mindset seeks tests that are self-referential. That is, the logic used to make a claim is the same as the logic used to test the claim…Such logic is self-referential in the sense that the individual state that the validity of his claim is that he knows it to be true…The consequences of defensive reasoning include escalating errors and self-fulfilling and self-sealing processes. This prevents testing not only the effectiveness of an outcome (single-loop learning) but also the assumptions behind the system (double-loop learning) (p. 114).

In contrast to defensive reasoning, Argyris advocates for productive reasoning, which he describes as being “In the service of learning, informed choice, and personal responsibility for the effectiveness of one’s actions” (p. 114). To execute productive reasoning requires analyzing actions and
Quantifying OD Results

results in a way that can be independently verified. Argyris writes that this type of objective verification allows practitioners to assure they are not “kidding themselves and others” (p. 114). Productive reasoning requires the type of third-party objectivity that cannot be attained from a person within the system they are evaluating. Hard data is particularly useful for this type of objective analysis, because numbers themselves are not swayed by sentiment (although of course the statistical presentation of those numbers can be influenced by the presenter/observers subjective perspectives).

Subjective Analysis vs. Objective Analysis

The severe limitations of subjective analysis are troubling when one looks at the responses to the anonymous survey question, “How do you know the difference between a successful intervention/engagement and an unsuccessful one?” Thirty-one respondents replied to this question; their answers were coded into three categories of: clearly subjective, unclear, and clearly objective (see Appendix D). Out of 31 responses to the question, 19 (61%) were clearly subjective anecdotal methods, making no reference to attainment of agreed upon objectives. These 19 subjective responses focused primarily on whether the client was happy or not, citing answers such as:

- “Client satisfaction.”
- “It can be difficult to tell. Mostly anecdotal from the client.”
- “If they re-engage me.”
- “The quality of change- it is palpable.”
- “Difficult to tell in the immediate I continue to keep a relationship with the organization over time to determine the effectiveness of the intervention.”
- “Happiness of the client, but not always. Sometimes I see a change in culture or behavior because of an engagement (even if it ends in the beginning stages and does not finish) and I often see those little changes as a success. Of course, the client probably wouldn't agree, but sometimes, I'm good with that.”
- “I expect some stake holders will feel much relief, some will feel much stress, and the majority will be confused or mostly unaware of the intervention.”
- “It is unsuccessful anytime the client is dissatisfied with my work.”
- “Success is in the eye of the beholder.”
- “By the level of energy when the last intervention is done.”
Quantifying OD Results

As shown in the examples above, the 19 very distinctly subjective analysis methods rely upon the subjective perceptions of the client, how the client or stakeholder feels, the level of energy, or palpable level of change. Feelings, perceptions, “levels of energy” are not units of measurement that can be verified by third parties. Even more concerning is that the subjective responses of the client are then interpreted through the consultant’s own subjective framework. Anecdotal conversation, as is so common in OD practitioner methodology of being “in partnership” with their client, is passed through two layers of subjective interpretation.

The subjective anecdotal nature of evaluative methods by current OD practitioners may help explain why many of the interview respondents reported that reaching project milestones is often substituted for tracking achievement of pre-determined hard data results. However, this approach of, ‘Did we do it?’ versus ‘How well did we do it?’ does not provide a means to engage in critical double-loop learning. As Schaffer and Thomson (1992) write, “Because of the absence of clear-cut beginnings and ends and an inability to link cause and effect, there is virtually no opportunity in activity-centered improvement programs to learn useful lessons and apply them to future programs” (p. 85). If OD is a learning discipline as espoused, we must work to align our theory-in-use. “Isn’t one of the key purposes of OD consultation to help human beings see where their position may lead to ineffective action?” (Argyris, 2005, p. 118).

Inductive Reasoning on the Lack of Objective Evaluation Practices

Argyris (1990) argues that the reason people utilize subjective defensive reasoning is to protect human beings and their systems from threat. Perhaps he is correct in implicating that there is a sense of security in not being held accountable by the objective measurement of hard data. The role of client/consultant “partnership” was heavily emphasized in the interviews as a critical factor to success. A side-effect of partnership is a loophole for accountability and perhaps consultants are not eager to eliminate that loophole.

There is also an inverse hypothesis, a conveyed sense of humbleness, which surfaced in a third of the interviews. Consultants seemed to believe there were great successes being accomplished during their
Quantifying OD Results

projects, but the credit for that success was not theirs to take. As one interviewee remarked, “My job is to make other people look good” (R#4). Many consultants spoke of approaching their clients with an “I’m here to help you attitude.” This attitude invariably sheds light on the source of the theme “Evaluation is the Clients’ Job,” assumingly due to a belief, held by practitioners, that the achievements and accomplishments should be credited to the client. Again, this may also contain a shadow side indicated by Argyris (1990) of the safety this lack of ownership also provides to the consultant.

There is a third hypothesis to explain the reason objective evaluations are not commonly performed in the OD field. Perhaps the ideological lineage of our epistemology leads to encouraging a type of evaluation practice, as represented in Campbell and Stanley’s (1963) 12 Threats to Validity, that simply is not a good fit for the practice of OD. As previously discussed in the Literature Review, during the formative years of the OD field, these 12 Threats to Validity were advocated as criteria to gauge the quality of evaluation research designs. The framework for this thinking was housed in the model of empirical testing.

The problem with evaluating group interventions empirically is that organizations and teams are ubiquitously unique. The conflicts, history, member status dynamics, friendships—all the incredibly intricate ingredients that comprise group dynamics are impossible to replicate to construct a control group. It is impossible to attain two identical groups. It stands to reason then, that early attempts at evaluation were highly prone to criticism when held against validity threats. And as Behaviorists would argue, humans are not prone to maintain doing what they know isn’t working; it is no surprise that the practice of evaluation in OD engagements atrophied. Not one of the 54 anonymous survey respondents claimed to utilize the 12 Threats to Validity as an evaluative guide.

What Can Be Done?

A recent publishing by Morris, Storberg-Walker and McMillian (2009) advocates developing an “OD-Intervention Metric System” with the use of an applied theory building method to examine work/life
balance issues. The study claims to adopt a systems view of human capital to evaluate the financial return on investments (ROI) of work/life interventions.

Disciplines such as Lean and Six Sigma are gaining popularity in business circles. These practices promote the use of metrics for Process Improvement methods. According to Arthur (2005):

Human performance, both individual and collective, is often measurable. OD practitioners may find themselves evaluating measures of productivity, error rates, absenteeism, training performance results or other assessments in search of clues to lapses in organizational performance or opportunities for improvement. The Six Sigma approach uses process variation as a key indicator of poor performance, an opportunity for improvement (when variation is diminishing), or as evidence for further examination or analysis. (p. 3)

It is interesting to compare the standard for “metrics” versus the standard for empirical evaluation. According to Becker, Huselid and Ulrich (2001) metrics possess credibility, can be believed by a reasonable person, have meaningfulness, are legitimate and accurate, reliable and valid, and possess strategic value for the end user. Typical metric categories consist of: volume, quantity, cost, income, time, quality, stakeholder reaction, rate, ratio, and categories or levels (Morris et al., 2009).

Missing from this list of metric requirements is the need for control groups and absolute elimination of internal and external validity threats, as is required in empirical testing. In contrast, “The foundation of a good metrics system is measurement of attributes and relationships that measure hypothesized causality as the goal for making decisions and assessing performance” (Morris et al., 2009, p. 425). Metrics allow much more freedom and flexibility and yet their credibility is not questioned. Quite the opposite: “In business, value must be measured through metrics” (p.445).

One example of this is the common Return on Investment (ROI) measurement. While the ROI is a simplistic instrument, it is widely accepted as a useful metric to measure business operations. Validity threats are not viewed as prohibitive from conducting evaluation. Certainly it does not have the precision of empirical science, yet it provides a useful evaluation framework. While an ROI evaluation is certainly
not a blanket answer, it is an indication that there may be other forms of evaluative methods than the OD community has not embraced.

**Quantifying the Qualitative**

As noted by the OD practitioners surveyed on the challenges of OD evaluations, the most frequently cited difficulty was quantitatively measuring “soft” criteria (Armenakis et al., 1976). Fortunately, new developments are being made in this area. In a study titled “The Role of Positivity and Connectivity in the Performance of Business Teams” researchers Losada and Heaphy (2004) coded the verbal communication of business teams, measuring the frequency of approving Positive (P) statements versus disapproving Negative (N) statements. This coding established a P/N ratio representing the frequency of positive verbal communication versus negative verbal communication.

Their research found that positivity acts as an amplifying or reinforcing feedback that expands behavior, opening possibilities for action, while negativity dampens behavior deviations. They found a statistically significant correlation between high performing teams and high P/N ratios, as well as low performing teams and low P/N ratios. Team performance—high, medium, or low—was established through an analysis of profitability (profit and loss statements), customer satisfaction (surveys) and 360-degree evaluations (assessments of the team members by superiors, peers, and subordinates.) According to the researchers of the study, high performing teams had a strong correlation rate of high P/N ratios, and demonstrated buoyancy, appreciation, and encouragement to their members during meetings. In contrast, low performing team meetings demonstrated lack of mutual support and enthusiasm, often creating an atmosphere of distrust and cynicism. The open encouraging space created by the high performing teams has the net effect of expanding the behaviors of its members to innovate and respond adaptively to continuously changing and challenging market and environment demands. Emotionally intelligent positive communication is often considered a “soft skill” and yet these researchers have been able to provide quantitative measures of its effects on business performance.
Quantifying OD Results

The work of Losada and Heaphy (2004) is an excellent example of the quantitative research void the OD field needs to fill to establish its credibility and relevancy to the business community. As Beer (1976) lamented, “OD suffers from an inability to demonstrate the relevance of its interventions. We desperately need research aimed at developing a technology for quantifying, in organizationally relevant terms (profits, services, good will, market competitiveness, innovations, etc.), the results of changes created” (p. 50).

Systemic Implications

As any systemic model will depict, organizations exist within a larger environment that impacts their well-being and ability to survive. Shifting environmental conditions are identified through an active feedback loop filtering between the organization and its environment that measures sales, buying trends, customer satisfaction, available resources, competition, etc. Organizations that choose to ignore these environmental influences, especially when that environment is shifting, tend to lack the ability to adapt and survive. As Pasmore (1988) writes, “How well the social and technical systems are designed with respect to the external environment determines to a large extent how effective the organization will be” (p. 1). The field of OD is not above or apart from these external environmental realities.

In practical terms, OD exists within a larger environment of the capitalistic business world. This environment is also populated with evolved forms of competition such as SixSigma, Lean, Change Management (all of which embrace the concept of metrics), and clients with reduced budgets. Like it or not, this current capitalist economy is where the livelihood of OD comes from; wouldn’t it make sense to adopt an evaluation practice that is compatible with the business world’s common language and orientation? “Too often among OD practitioners one finds a disregard or even hostility toward the profit motive in business organizations….Profit is not our only goal. It is far too narrow a concept to use in defining our work and our mission, but we must include it in our work and be happy to pursue this as well as other goals” (Petrella, 2005, p. 68). Yet, the majority of evaluations respondent in both the interviews
and anonymous survey disregard whether or not the engagement affected the organization’s bottom line. The OD field’s lack of adaptation to the marketplace’s orientation indicates there is a critical feedback loop in our own systemic model that is dangerously underdeveloped. Bradford and Burke (2005) go so far as to write, “There is a reticence among OD consultants to look at the dysfunctional effects of their interventions” (p. 207).

Implications for the Field of OD

Based on the data and literature research, it is clear that the field of OD can strengthen its presence in several ways. First, during client engagements, determining a mutually agreed upon goal and means of distinguishing and measuring traction, is essential before beginning the project. Ideally, this would be a metric that can be measured objectively, drawing from the metric categories listed above or inspired by a SMART goal: specific, measurable, attainable, relevant and time-bound. This is not to say that objectives cannot change over the course of the project, as new information comes to light. As many consultants are fond of saying, “the presenting problem is never the real issue or what you work on” but as objectives change, success metrics can be re-determined and re-contracted.

This shift in contracting behavior requires spending time to define qualitative aspirations. Often success goals are established such as “increased communication,” “reduced conflict,” or “streamlined meetings.” However, a second step is needed to spend time considering how the consultant and client will know when communication has increased, how much conflict has been reduced, and what behaviors will qualify as a streamlined meeting. “In a very basic way, every project is about reconciling the fundamental polarity between the world of What-Is-Needed and the world of What-Can-Be-Built” (Conklin, 2006, p. 16). Objective parameters by which to measure these larger goals need to be specified. What-is-Needed might be increased diversity, energized innovation, or an improved level of employee engagement - but these are aspirations that have no ceiling for achievement. A conversation to balance these what-is-needed items must be accompanied with an honest what-can-be-built perspective. For “increased diversity” a goal
Quantifying OD Results

might be set of an awareness training program and a 7% change in hiring or attrition statistics. “Energized innovation” might be objectively tracked by the number of employee contributions, engineering design submissions, item-to-market ratios, or number of new products released to market. “Employee engagement” might track attrition levels, employee satisfaction surveys, or numbers of colleague referrals. Achieving these identified what-can-be-built targets will not mean that the larger what-is-needed issues will have been “solved” but it does provide the consultant a means to objectively reflect on their methodologies’ success indicators. Ideally these identified metrics will provide the client a chance to evaluate how their financial investment in the OD project is impacting their organization’s profitability.

The problem with large scale activity-based change efforts, where a what-can-be-done objective parameter is not identified, is that connecting the activities to the bottom line is complicated by the fact that a vast array of activities are launched simultaneously. As Schaffer & Thomson (1992) write, “This is like researching a cure for a disease by giving a group of patients ten different new drugs at the same time” (p. 83). In contrast with activity-centered approaches, “the results-driven path stakes out specific targets and matches resources, tools, and action plans to the requirements of reaching those targets. As a consequence, managers know what they are trying to achieve, how and when it should be done, and how it can be evaluated” (Schaffer & Thomson, 1992, p. 82). As the interviews showed, existing metrics within organizations is an excellent place to begin looking for a measurement metric.
Conclusion

The philosophical humanistic underpinnings of OD such as emphasis on self-awareness, communication, authenticity, inclusion and collaboration bring a perspective into the workplace that enriches lives and serves to counter-balance the profit-centered approach of the modern economy. However, if we cannot adapt to translate these contributions into a language the current external environments understands and values, the reach and impact of OD will decline. As Bradford and Burke (2005) write, “The consequences are that they [OD practitioners] will not be at the table where significant change decisions are made. Instead, at that table will be representatives from the major consulting firms. Only the crumbs will be left for the OD practitioner” (p. 213). OD practitioners need to pay attention to this risk and bring it into our collective narrative.

The OD field’s lack of adaptation to the marketplace’s orientation indicates there is a critical feedback loop in our own systemic model that is dangerously underdeveloped. We are not listening and adapting to the larger community upon whose resources we rely for survival. We need to figure out how to capture the qualitative shifts that occur during an OD intervention and translate them into quantitative examples. It is self-sabotage to remain elitist in our language and our values. The client and the greater community’s values, such as profit, must enter into our thinking and design if this field is to stay relevant and not become extinct.

The feedback loop needs to be repaired and strengthened. This discussion needs to enter our conferences and professional gatherings. Contemporary evaluation instruments need to enter our research focus. Business acumen and results evaluation need to be incorporated into graduate programs’ curriculum, and finally, as practitioners, we need to strive for dedicated personal accountability during our own engagements.
References


Quantifying OD Results


Appendix A: Interview Questions

Introductory questions:

- How long have you been an OD/OE practitioner?
- Is the majority of your work internal/external?
- Where do you typically enter your client organizations? At what level? What is the typical job position/level of your client contacts?

Scoping/Contracting phase:

- Do you discuss expectations for the success of your intervention during the contracting stage?
- Does this discussion include “hard data” expectations?
- If no, how are expectations discussed?

Post-or current-Intervention questions:

- Do you perform an evaluation of the impact of your OD intervention when it is completed?
- Do you use evaluation to steer your intervention mid-course?
- How specifically do you evaluate the results of your intervention?
- Have you ever evaluated your results against Campbell & Stanley (1963) twelve threats to validity? Are you familiar with it?
- Have you ever evaluated your results in an ROI framework? Are you familiar with them?
- Have your evaluations changed or morphed over time? How?

Reflective questions:

- What are the effects of the evaluation on your self-perception? Do you think that your evaluation method affects the way your client perceives you? If so, how?
- Do you use your evaluation results when you speak of yourself during the contracting phase?
- Do you think you gain more credibility with potential or current clients because of your evaluation metrics? Why do you think yes or no? Do you believe you have greater access to top executives based on your ability to discuss how your intervention affects the bottom line of the organization? Why do you think yes or no?
1. Please rank the following professional titles you most utilize in professional circumstances? With 1 being the most often used to 9 being the least often used. Or select N/A if you never utilize the title.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>N/A</th>
</tr>
</thead>
</table>

2. How long have you been working as a practitioner in the above field(s)?

- [ ] 0-3 years
- [ ] 4-5 years
- [ ] 6-10 years
- [ ] 11-15 years
- [ ] 16-20 years
- [ ] 21-25 years
- [ ] 26-30 years
- [ ] 31+ years

3. Is the majority of your work:
Quantifying OD Results

☐ Internal?
☐ External?
☐ Combo?

4. What is the typical job position/level of your client contacts who hire you to work with the organization or team/department?

(Check all that apply)

☐ Executive Team
☐ VP
☐ Director
☐ Manager
☐ Line/Staff

5. What level in the organization do your projects typically work with?

(Check all that apply)

☐ Executive Team
☐ VP
☐ Director
☐ Manager
☐ Line/Staff

6. On average... (please circle)

Do you discuss expectations for the success of your intervention during the contracting stage?
Never Rarely Sometimes Frequently Always

Does this discussion include “hard data” expectations?
Never Rarely Sometimes Frequently Always

7. Please check the boxes of all the methods you utilize when evaluating the results of your projects:

I do not evaluate the results of my projects
Quantifying OD Results

- Surveys and/or questionnaires
- Anecdotal Conversations
- Common Business Metrics (i.e. profit, loss, inventory, attrition, time to market...)
- Another form of metrics (possibly created by you, i.e. number of employee errors, conflicts, process time)
- Campbell & Stanley (1963) Twelve Threats to Validity
- ROI (Return on Investment) analysis

8. Please rank, in order of importance, how you establish credibility when you begin contracting with a new client. 1 represents most important, 6 least important.

   1 2 3 4 5 6

   With financial figures or metrics of previous engagements
   Word of mouth/reputation/referrals
   By demonstrating business acumen in conversation
   By asking the right questions
   By discussing previous anecdotal successes, i.e. design or process
   By forecasting probability of success for the proposed project

9. How do you know the different between a successful intervention/engagement and an unsuccessful one?

10. What encourages or discourages you from performing an evaluation at the end of an engagement?
## Appendix D: Coded Survey Response answers to Question #9

<table>
<thead>
<tr>
<th>Q9. How do you know the difference between a successful intervention/engagement and an unsuccessful one?</th>
<th>Clearly subjective</th>
<th>Unclear</th>
<th>Clearly objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/24/11 12:22AM: The client and I define what will be considered successful up front (e.g., improved communication among team members). If the criteria is met then it was a successful project.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 12:26AM: Whether the customer is happy or not.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 12:32AM: clients’ responses &amp; feedback</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 12:34AM: By the commitment of the leader around what it is that he/she wants to cause.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 12:36AM: success will realize the projected outcomes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 12:36AM: 1) consistent and sustainable follow through of plans upon completion of business meeting 2) Candid and transparent dialogue with business manager/operator 3) Full participation and engagement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 1:26AM: Situational- Did you meet the agreed upon objectives for the engagement?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 1:48AM: good evaluations, call back for more work, attainment of stated goals, constant communication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 3:09AM: Client’s definition combined with internal department metrics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 3:12AM: The clients expectations and outcomes have been met. Occasionally there will be a report with a clear ROI business case.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 3:35AM: It can be difficult to tell. Mostly, anecdotal from the client. Also, if they re-engage me.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 4:27AM: Client satisfaction. Follow-up later. Referrals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 4:28AM: In discussions with the client manager. Also, for #8, I would say that Listening to what the client has to say is actually #1 for me in establishing my credibility.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 4:57AM: Client is empowered and equipped to build on their success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 6:06AM: the quality of change - it is palpable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 7:35AM: at the beginning we have outlined the outcomes for a successful project - whether or not we've met these outcomes.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 7:42AM: The client opens their organization to me if it's successful. The client says thank you and good bye, you've been a great help if unsuccessful.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 4:34PM: Difficult to tell in the immediate. I continue to keep a relationship with the organization over time to determine the effectiveness of the intervention.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24/11 7:59PM: Whether the metrics identified in contracting before and during the engagement are met or exceeded determines the degree of success.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/11 4:15AM: Client perception, team competency development, impacted agent feedback on process</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/11 4:29AM: Happiness of the client, but not always. Sometimes I see a change in culture or behavior because of an engagement (even if it ends in the beginning stages and does not finish) and I often see those little changes as a success. Of course, the client probably wouldn't agree, but sometimes, I'm good with that.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/11 5:36AM: If we have met the stated objectives of the engagement.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/11 5:44PM: Depends on the time horizon. Often I find that the grandest long term successes feel the most chaotic. When our engagement or contribution is &quot;over&quot; people generally tell you how they feel it went. If it was a big and complex change which spans a decade, and we did a 3 month intervention, I expect that some stakeholders will feel much relief, some will feel much stress, and the majority will be confused or mostly unaware of the intervention.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/11 6:28PM: It is successful when the client feels satisfied when I leave, even if we didn't achieve the original outcome. It is unsuccessful anytime the client is dissatisfied with my work. Even if I achieved contracted objectives, if the client does not feel it was successful there was a disconnect somewhere in the process.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/11 7:04PM: When people can see that they’ve moved toward where they want to go. The metrics/indicators for that vary from place to place. Success is in the eye of the beholder.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/11 9:25PM: By the completion of deliverables, learning objectives, capacity building. A key effort of mine is to ensure that I build the capacity of any person I work with.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/26/11 7:41PM: It depends on the agreed upon goal established at the beginning of the project - and it is a recap conversation toward the end of the project.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/28/11 2:25AM: Discussions with the client afterwards, recognizing that an organization has adopted new ways of doing things, the ideas persist and they tell me they are pleased.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/28/11 4:05AM: by the level of energy when the last intervention is done, by debrief with client manager and by status quo on the results in forms of action planning and other tangible outcomes.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/28/11 11:03PM: Was the project implemented? Does the client think the project was a success? Were their metrics? Will/can I use the client as a reference?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/6/11 4:52PM: Business results. If the end state is not better than the beginning state, then no change has occurred. If that's the case, the intervention was unsuccessful.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>