The Use and Validation of Preintervention Diagnostic Tools in Organizational Behavior Management

DOUGLAS A. JOHNSON, SARAH E. CASELLA, HEATHER McGEE, and SIN CHIEN LEE
Western Michigan University, Kalamazoo, Michigan, USA

Functional analyses and assessments have become the accepted gold standards in many applied areas of behavior analysis, including Organizational Behavior Management (OBM), in recent years. Despite their acceptance, OBM data on such tools have been largely absent. There are several assessment alternatives in OBM (e.g., ABC, PIC/NIC, Behavior Engineering Model), but two assessment approaches most common in the Journal of Organizational Behavior Management are Behavioral Systems Analysis (BSA) and the Performance Diagnostic Checklist (PDC). To date there have been no comprehensive reviews of BSA or the PDC from which one might draw ideas for application, research, and the advancement of the field. The goals of this paper are to (a) provide a review of the BSA and PDC literature within the Journal of Organizational Behavior Management (JOBM), (b) discuss the implications of the results of the review, and (c) provide suggestions for future research utilizing BSA and the PDC.

KEYWORDS behavioral systems analysis (BSA), Performance Diagnostic Checklist (PDC), preintervention analysis

During the past several decades, functional analysis and functional assessment have become the accepted standard in many applied areas of behavior analysis (Bailey & Burch, 2002; Iwata, Dorsey, Slifer, Bauman, & Richman, 1982). This paradigm shift has moved the field to first assess the variables maintaining behavior or performance excesses/deficits and then use the
results of that assessment to determine the most appropriate intervention(s). Organizational behavior management (OBM) is one such applied area in which authors have called for assessments being conducted either a priori or concurrently when designing performance improvement interventions (Austin, Carr, & Agnew, 1999).

As a method for deciding whether this call has been answered, we reviewed all issues of the *Journal of Organizational Behavior Management* (*JOBM*) from 1977 through 2011 (book reviews were not included), Volumes 1–31, in order to determine the current state of the use of assessment in OBM studies. We initially examined articles for at least one of the following terms: assessment, systems analysis, systems thinking, processing system, receiving system, process maps, relationship map, total performance system, performance matrix, super system, organizational scan, behavior engineering model, human performance system, PIC/NIC, behavior engineering model, ABC analysis, or Performance Diagnostic Checklist. This list of terms was first generated based on our knowledge and expertise in the field. The list was expanded by conducting an informal review of OBM books and articles to identify additional terms used when describing assessments in OBM. Our review suggested that there are several assessment alternatives in OBM, but two assessment approaches have been particularly visible (i.e., appear multiple times when search terms such as assessment or diagnostic are used) in *JOBM*, namely, behavioral systems analysis (BSA) and the Performance Diagnostic Checklist (PDC). These two assessment approaches are described in more detail next.

**BSA**

BSA is a methodology for improving performance that blends applied behavior analysis (the application of the science of human behavior to the prediction and control of behaviors that are considered socially important; Bailey & Burch, 2002) and general systems theory (an approach to understanding organizational systems by examining the relationships between parts of the system, as well as the relationships between the system and the external environment; Ackoff & Emory, 1972; Bertalanffy, 1950, 1968). With its focus on both the system and the person, BSA offers a framework for understanding how organizations work and how to make them work better and has evolved through the work of many pioneers in the areas of behavior analysis and performance improvement (Brethower, 1982; Daniels & Daniels, 2004; Gilbert, 1978; M. E. Malott, 2003; R. W. Malott, 1974; Rummler, 2004; Rummler & Brache, 2013; Skinner, 1953).

BSA has a long history (Brethower, 1972, 1982; Connellan, 1978; Harshbarger & Maley, 1974; Morasky, 1982) and is a very visible part of OBM, as evidenced by the special issues dedicated to BSA in Volume 29 of
The field of BSA has continued to grow and develop an elaborate and impressive series of maps, charts, and other tools intended to be populated via observations, questionnaires, and guided interviewing to aid in the pinpointing of performance gaps in complex environments as well as in guiding the subsequent change efforts that result from the identification of those gaps.

Behavioral systems analyses typically result in multilevel solutions that may include performance management interventions, process design, automation, policy changes, resource allocation, strategy development and/or realignment, the development of incentive systems, organizational restructuring, performance-based training systems, and managing-the-manager initiatives (Diener, McGee, & Miguel, 2009; McGee, 2007). The value of BSA is that it considers identifying, planning, and managing the factors that significantly impact individual and organizational performance, which could impact the sustainability of performance change efforts. In fact, Redmon (1991) called for the explicit use of systems analysis in all published studies as a means of encouraging program adoption by organizations in which OBM interventions are used.

THE PDC

The PDC was developed based on the results of Austin (1996), a study that provided consultants and managers with a series of performance problems and asked them to talk aloud as they tried to solve the performance issues. The results indicated that the participants who were most successful in solving the performance problems asked questions related to the current state of antecedents, equipment and processes, knowledge and skills, and consequences. Austin (2000) developed a series of diagnostic questions around these four areas, referred to as the PDC. The answers to these questions guide the selection of interventions aimed at improving performance. Multiple studies have asserted that the PDC is the most commonly used assessment tool in OBM today (Fante, Gravina, & Austin, 2007; Fante, Gravina, Betz, & Austin, 2010).

COMPARING BSA AND THE PDC

Both BSA and the PDC involve the use of tools to diagnose the functioning of an organization and pinpoint opportunities for improvement. Both approaches rely in part on informant assessment to infer functional relations, though proponents of both would recommend that direct observation and data be used to confirm results of the informant assessment as much as possible. Thus, there is a fair amount of overlap between these assessment
approaches. Both approaches attempt to pinpoint disconnects in work processes, identify necessary skills and resources for performing work, and determine the relation between performance and the environment.

There are several noteworthy differences as well. BSA tends to assign a greater emphasis to processes and relations (Rummler & Brache, 2013), whereas the PDC places a greater emphasis on environmental antecedents and consequences (Austin, 2000). There are also formal differences between these assessment approaches. Unlike the single checklist that mainly targets the job or performer level of performance involved in PDC assessments, BSA tends to involve multiple tools to diagnose goals, structure, and management at multiple levels (typically the organization, function, process, and job or performer). Thus, BSA is more likely to provide a more thorough and multifaceted picture of an organization, but such an assessment will likely come at the price of a greater upfront time investment compared with the PDC. Many of the BSA tools involve the creation of highly visual depictions of the organization (such as flowchart tools). Representing structures and workflow visually may lend itself to easier communication between the consultant and client or facilitate the consultant’s pinpointing of redundancies or gaps. However, the many maps and diagrams may prove cumbersome when contrasted with the PDC’s simple yes/no checklist format, confined to a single page. Nonetheless, such relative strengths and weaknesses of these assessment tools remains an open empirical question.

STATEMENT OF THE PROBLEM

During our review of the BSA and PDC assessment approaches in JOB, a noticeable problem began to emerge. Namely, there appeared to be little evidence to suggest empirical validation of either approach. To support this observation, we reviewed the BSA and PDC articles more carefully. We only considered BSA and PDC articles that included a substantial discussion of the tools/approaches (a description of how it was used in experimental studies or comprising, at least in part, the topic of the article for theoretical or conceptual articles). We also classified the articles according to the level of empirical comparison. We utilized the following categories for classification: theoretical or conceptual, minimal empirical data, comparison with control, and comparison with alternative.

The category of theoretical or conceptual was used for articles that were based purely at the level of discussion or theory. Articles reporting hypothetical data were included in this category. Review articles were also included in this category if they did not present previously unpublished data sets. Discussion articles would clearly fit within this category, although some research articles could also be candidates for this category. Research articles that presented data, even though the intervention was not driven by BSA,
were included in this category and indicated in Table 1 (no similar designations were necessary for the PDC articles). For example, an article that included a discussion of BSA, but for which the interventions were not chosen based on the use of BSA tools or processes (or at least it was not made clear in the Method section that this was the case), were included in this category. It is possible that the interventions were chosen based on a systems analysis, but one cannot know this without a description of it. This category was developed to capture any substantial articles regarding the diagnostic tools of interest that may not have involved empirical data. Such a category may prove useful in highlighting the degree to which the field simply discusses diagnostic tools compared to the degree to which the field actually utilizes or examines such tools.

TABLE 1 Classification of Behavioral Systems Analysis Articles Published in Volumes 1–31 of Journal of Organizational Behavior Management

<table>
<thead>
<tr>
<th>Author(s) and year</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abernathy (2008)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Abernathy (2009)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Alavosius et al. (2009)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Austin et al. (1999)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Berglund &amp; Ludwig (2009)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Brethower (2000)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Brethower &amp; Wittkopp (1988)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Clayton et al. (1997)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Diener et al. (2009)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Frederiksen et al. (1985)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Gikalov et al. (1997)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Hayes et al. (2009)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Houmanfar et al. (2009)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Huberman &amp; O’Brien (1999)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Hyten (2009)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Krapfl (1982)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Krapfl et al. (2009)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Kriesen (2011)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Langeland et al. (1998)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>R. W. Malott &amp; Garcia (1988)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Mawhinney (2009)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>McSween &amp; Matthews (2005)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Mihalic &amp; Ludwig (2009)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Riley &amp; Frederiksen (1984)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Sandaker (2009)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Sasson et al. (2006)</td>
<td>Comparison with control; comparison with alternative</td>
</tr>
<tr>
<td>Sasson &amp; Austin (2002)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Smith &amp; Chase (1990)</td>
<td>Theoretical or conceptual</td>
</tr>
<tr>
<td>Sulzer-Azaroff et al. (1993)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Tosti &amp; Herbst (2009)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Williams et al. (2002)</td>
<td>Theoretical or conceptual</td>
</tr>
</tbody>
</table>

aData-based article, but data presented did not involve behavioral systems analysis.

bUtilized a 2 × 2 factorial design that allowed for multiple types of comparisons to be made.
The category of *minimal empirical data* was used for articles that presented empirical evidence (i.e., actually used BSA or the PDC) but utilized designs that prevented the establishment of cause-and-effect relations regarding enhancements to the interventions resulting from a particular preintervention diagnostic tool. Examples would include case studies and AB designs.

The category of *comparison with control* was used for articles presenting empirical evidence that helps establish that interventions selected through the use of BSA or the PDC are better than (or worse than) no active intervention (or no change to existing interventions or initiatives).

The category of *comparison with alternative* was used for articles that compared interventions selected through the use of BSA or the PDC to another active intervention selected by another type of analysis (or no prior analysis).

### SUMMARY OF JOB M ASSESSMENT LITERATURE

Our attempts to classify assessment tools in *JOB M* indicated that out of 31 BSA articles reviewed, approximately 71% of BSA articles were theoretical or conceptual in nature, 26% of BSA articles provided minimal empirical data, 3% of BSA articles involved comparisons against a control, and only one BSA article involved comparisons against an alternative (see Table 1). It is important to note that the article by Sasson, Alvero, and Austin (2006) was classified as both a control and an alternative approach condition because it utilized an experimental design that allowed for multiple types of comparisons. Specifically, a $2 \times 2$ design was used to compare two types of process variables as well as compare intervention and no intervention conditions.

Results of the final review also indicated that out of eight PDC articles reviewed, approximately 12.5% of PDC articles were theoretical or conceptual in nature, 87.5% of PDC articles provided minimal empirical data, and none of the PDC articles involved comparisons with a control or comparisons with an alternative (see Table 2).

### TABLE 2 Classification of Performance Diagnostic Checklist Articles Published in Volumes 20–31 of *Journal of Organizational Behavior Management*

<table>
<thead>
<tr>
<th>Author(s) and year</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amigo et al. (2008)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Doll et al. (2007)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Eikenhout &amp; Austin (2005)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Gravina et al. (2008)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Pampino, MacDonald, et al. (2003)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Rodriguez et al. (2005)</td>
<td>Minimal empirical data</td>
</tr>
<tr>
<td>Weatherly &amp; Malott (2008)</td>
<td>Theoretical or conceptual</td>
</tr>
</tbody>
</table>
IMPLICATIONS FOR ASSESSMENT IN OBM

The publication of the first functional analysis (Iwata et al., 1982) led to a paradigm shift in applied behavior analysis toward preintervention selection assessments and analyses of the variables maintaining/suppressing behavior. However, even in the year 2000, OBM researchers and practitioners were still noting that assessments were not reported in the majority of published OBM studies (Austin, 2000; Austin et al., 1999). Since that time, it would appear that the use of a priori or concurrent assessments has increased, as evidenced by the number of articles using two common approaches to organizational assessment: BSA and the PDC. Despite their increasing prevalence, and despite calls for action by experts within the field, preintervention assessments are not explicitly incorporated into the majority of OBM articles. From 2000 to 2011, less than 15% of published JOBM articles made more than a brief mention of BSA or PDC tools, despite the fact that this time period included multiple special issues dedicated to these topics. A cursory look through the remaining 85% of articles suggests that it is not the case that other assessment tools are being utilized. Although it is not appropriate that assessment tools be mentioned in every article, the data strongly suggest that the published use of these tools remains a rarity. Despite assertions regarding the popularity of the PDC (Fante et al., 2007, 2010), the use of this tool seems to be limited and possibly in decline. During the last 3 years of this review (2009–2011), no published articles used the PDC.

As we suggested earlier, there is little evidence to suggest empirical validation of either approach. In fact, although 26% of the BSA studies and 87.5% of the PDC studies provided evidence that the tools were used to guide the selection of interventions, only one BSA study and zero PDC studies actually attempted to empirically validate the assessment approach through experimental manipulations. Although a single BSA study nominally analyzed systems variables (Sasson et al., 2006), this study only looked at simple variations of a process (deliver materials via manual or electronic means) and may not well represent the lengthy processes seen in applied settings or the typical in-depth analyses that result from examining such processes. Thus, it could be argued that there are no true exemplars of BSA or PDC experimentation in the existing literature. The lack of empirical validation may be a contributing factor to the standard use of these tools in published accounts of organizational interventions. There are many possible reasons for the lack of empirical validation of OBM assessment approaches.

Environmental Complexity

One potential reason for the lack of directly comparing BSA or the PDC to either a control condition or an alternative assessment approach is that these assessments occur within complex environments. In complex environments,
it is difficult to control extraneous variables well enough to determine functional relationships. Researchers could mitigate this by first conducting research in the lab to determine whether functional relationships exist and then following up with field studies to assess the generality of the findings. However, recreating a sufficiently complex environment in the lab is difficult, if not impossible, especially for the purposes of testing a systems analysis. Even if researchers could create a sufficiently complex environment, it is still possible that an investigator–environment interaction would create an experimental confound.

Whereas functional analyses used in other areas of behavior analysis rely on the direct observation of behaviors under a variety of tightly controlled conditions, both BSA and the PDC (and many other assessment approaches) require investigator interaction with a complex environment throughout the assessment process. This may lead to an unintentional altering of the environment even before an intervention is implemented. For example, suppose a researcher wanted to compare BSA against the PDC to determine whether they yield different recommendations for interventions. Both approaches utilize informant assessments. If the researcher conducted the assessments within the same organization, the act of interviewing an organizational leader during one assessment would have a strong potential to affect the responses the same leader would give if interviewed again using the other assessment. If this were to occur, it would contaminate the information gathered during the alternative assessment. If the researcher tried to control for this by interviewing a different individual, even if that individual held a position at the same level of the organizational hierarchy, there is no guarantee that the second individual would have the same knowledge of the organization as the first individual. This suggests the need to compare across sites or conditions instead.

Suitability of Comparison Sites or Conditions

If significant confounds are likely in a study that compares two assessments in the same setting, another option would be to try to either create or find two very similar settings in which to conduct the assessments. As previously stated, recreating an environment that is complex enough to warrant a systemic analysis poses difficulties. Complex performance problems would need to be created. Because the PDC and BSA both involve informant assessments, these environments would also need “employers” and “employees” who possess information (both facts and opinions) about the performance issues at hand. This might not be feasible or practical.

An alternative to creating the organization in a lab setting is to find two similar settings. Though this may still be difficult, it is at least feasible. Possible options would be to look for either two similar sites (e.g., two different manufacturing plants) of the same organization or perhaps two different
organizations (e.g., two different franchises) under the same parent company (e.g., a fast food restaurant). Of course, no two sites or organizations are the same, which could potentially impact the results if the investigator is attempting to compare two different assessments or different components of the same assessment. However, this might be an acceptable approach if one is comparing the results of one assessment type against a control condition, particularly if a multiple-baseline-across-settings design is used in which measures are held constant across the sites.

Cost and Time

Two other potential reasons for the lack of empirical investigations of OBM assessments are the cost and the time required to conduct these assessments. This is of particular concern with BSA, which can take several weeks or even months (Rummler, 2004). Although practitioners get paid to conduct their assessments, the previously mentioned issues might preclude an experimental analysis of the validity of the assessment. For academic researchers, the potential cost and time requirements to find or create the appropriate organizational environments and carry out the assessments might outweigh the perceived value of conducting the investigations. With both practitioners and academic researchers, there is a benefit in that extended analyses are likely to result in more complete information, allowing for a more sophisticated and individually tailored intervention. However, for both types of researchers there is a delay in creating organizational change. This may be a delay that impatient clients, particularly those not persuaded by scientific rigor, will not tolerate unless measurably superior results can be shown or it can be proved that a lack of detailed upfront assessments may risk costly errors or interfere with future assessment results and interventions. On a related note, empirical evidence regarding the suboptimization principle could be collected. This principle suggests that optimizing the effectiveness of one aspect of an organization without consideration of other aspects of the organization may result in a detrimental impact to the organization as a whole (M. E. Malott, 2003). If such a principle could be empirically demonstrated, it could serve as a justification to the client for the extra time and effort involved, particularly with potentially lengthy assessments such as those required by BSA. Without such an empirical demonstration, it is possible that clients may prefer something like the PDC because it is a quick and simple checklist, often based on a single interview. However, BSA tends to involve multiple tools to diagnose organizational structure and workflow at multiple levels. This often necessitates information being collected from many sources, potentially reducing the bias inherent in an interview with a single individual. Thus, it is possible that BSA would provide a more objective picture of an organization’s processes compared to the PDC. In other words, when using a PDC checklist, one may risk overlooking the root cause of disconnects that
Preintervention Diagnostic Tools

may otherwise be noticeable if a BSA tool were used. Of course, this remains pure speculation and reiterates the need for more empirical investigations.

Suggestions for Future Research

Based on our review of the literature, we would suggest that OBM has yet to produce studies to validate common assessment procedures or determine optimal assessment approaches. The majority of the empirical evidence currently available appears to comprise studies that use AB designs and one-group pretest–posttest designs (classified as minimal empirical data in both tables). Because these designs lack elements of replication or verification, such studies do not effectively establish cause-and-effect relations (Cooper, Heron, & Heward, 2007; Shadish, Cook, & Campbell, 2002). The plurality of anecdotes and case studies, although useful in describing how BSA and the PDC are used in organizational settings, are not sufficient forms of empirical evidence for the use of preintervention diagnostic tools. OBM needs to extend its research agenda further to answer (a) whether preintervention assessments enhance the selection of interventions by suggesting those that have the greatest positive effect on the performance of the individual and the organization and (b) whether that added benefit is worth the added cost. Thus, we offer some suggestions for future research.

Experimental evidence involving BSA or the PDC compared against a control condition

As stated previously, two possible options would be to look for two similar sites (e.g., two different manufacturing plants) of the same organization or perhaps two different organizations (e.g., two different franchises) under the same parent company (e.g., a fast food restaurant) and compare the results of either BSA or the PDC against a control condition (no assessment conducted prior to choosing interventions). Critical performance measures could be defined prior to baseline data collection using the same measures across sites. A multiple-baseline-across-settings design could then be used to assess whether interventions suggested by either BSA or the PDC impacted those predefined measures over interventions chosen without the guidance of an a priori assessment.

Experimental evidence involving BSA or the PDC compared against an alternative approach condition

Although the conditions would not be perfect, the previously described scenario for comparing BSA or the PDC against a control condition could also be used to compare one assessment approach against an alternative approach, though the design would need to be altered to a between-groups
or mixed between-groups and single-case multiple-baseline design. If the goal of such a study is to simply assess whether the different approaches yield similar suggestions for interventions, it might be possible to compare two sites, provided they are similar enough that one could assume the same causal factors surrounding the performance issues (e.g., two manufacturing plants with the same physical layouts, processes, measures, and employer/employee characteristics). Alternatively, researchers could instead look for different individuals within the same setting, again provided that the researchers can demonstrate that each individual has similar knowledge of organizational performance factors (e.g., two different plant managers or supervisors from the same plant but possibly different shifts). Researchers could conduct a study comparing BSA assessment tools or the PDC checklist to other organizational analysis strategies (e.g., PIC/NIC analysis, various descriptive assessments, experimental analyses, intuitive guessing). There are multiple measures by which these strategies can be compared: differences in terms of the types of interventions that are selected, total time to select and implement interventions, success of interventions driven by various analyses, and so on.

The interventions identified by various assessment procedures could be compared in a number of ways, including stimuli typically targeted for intervention (antecedents or consequences), individuals targeted for intervention (management or subordinates), complexity of proposed intervention (single component or package), longevity of proposed interventions (short-term or long-term impact), performance aspect targeted for intervention (behavior or result), level of performance targeted for intervention (e.g., performer, process, function), and other relevant considerations. Reviews of the publication history of *JOBM* (Balcazar, Shupert, Daniels, Mawhinney, & Hopkins, 1989; Nolan, Jarema, & Austin, 1999; VanStelle, S. E., Vicars, S. M., Harr, V., Miguel, C. F., Koerber, J. L. et al., 2012) could provide the relevant intervention comparison criteria. If it were discovered that the type of proposed intervention tends to differ as the type of assessment differs, this would be a valuable contribution to the literature and would suggest that the selection of a diagnostic tool may depend on the goals of the diagnostician.

**EXPERIMENTAL EVIDENCE INVOLVING THE COMPARISON OF ONE OR MORE BSA TOOLS OR THE COMPARISON OF ONE OR MORE PDC COMPONENTS**

Several dismantling studies could be conducted to determine better methods for analysis, preferably methods that are more time efficient when the unneeded excess is trimmed away. For instance, one could examine whether the mission statement needs to be considered and communicated (a standard practice in BSA but not in PDC) and, if so, to how many levels. Also, many systems tools rely on visual representations of organizational structures and processes (e.g., total performance system or process map). How such visual
Preintervention Diagnostic Tools

representations compare to text-based job aids in terms of both time to create these assessment tools and their value in pinpointing disconnects should be investigated. Researchers could also investigate whether all possible levels (i.e., macrosystem, organization, process, task, behavior, management) of analysis are necessary to conduct an assessment for a successful intervention compared to one or two levels of analysis. Researchers could also empirically analyze the various BSA tools (i.e., total performance system, process map, relationship map) to determine which tools are essential for identifying the appropriate intervention(s) and which tools are excess. The importance of each of the individual components of the PDC checklist could also be experimentally analyzed. For example, the PDC checklist contains a survey item asking whether employees can see the effects of performance. Whether this particular item (or any other component of the checklist) adds value to intervention selection could be examined through various dismantling strategies. It could also be examined whether the addition of any new items to the checklist enhances intervention selection and implementation.

EXPERIMENTAL INVESTIGATION OF POSSIBLE INVESTIGATOR EFFECTS

As previously mentioned, BSA and the PDC both rely to some extent on informant assessment. Informant assessments rely on interviews, surveys, or other methods of gathering verbal reports of current performance, performance gaps, and potential causes of those gaps from key performance stakeholders. Depending on the performance(s) in question, these stakeholders could be leaders, managers, supervisors, customers, suppliers, and/or employees. We have argued that comparing different assessment tools in the same setting by interviewing the same individuals could be difficult if the use of one assessment results in an investigator–environment interaction that affects the results provided on another assessment. However, this investigator–environment interaction effect has not been experimentally validated. Researchers could experimentally investigate whether such an effect exists. Again, this would necessitate the identification of comparable sites in advance (such as four separate franchises with similar policies, problems, and employees). Researchers could potentially utilize a multiple-baseline design across settings (taking care to ensure that the target behavior(s) is the same and the settings are similar): assess one site using BSA, a second site using the PDC, a third site using BSA and then the PDC, and a fourth site using the PDC and then BSA. Sites that only use either the PDC or BSA could serve as constant series control. This research idea could answer many questions about the relative merits of BSA versus PDC and the potential contaminations/enhancements resulting from the cumulative effect of multiple assessment tools for a single site. Although this would obviously be a difficult study, it could be completed as a series of studies to ensure functional control.
Alternatively, a single site could be assessed by different independent investigators. For example, two investigators could both use BSA or the PDC to diagnose the organization’s functioning and propose interaction. The degree to which these independent assessments match could be very important in determining the reliability of various assessment approaches.

Concluding Comments

As with any critical review, there are some potential limitations of the present discussion. It is possible that the search terms used were not broad enough and as a result did not allow for certain assessment articles to be included in the review process. Another limitation is that we only considered articles from *JOBM*. It is possible that other publications outside of the reviewed journal contain additional examples of BSA or PDC experiments. We specifically chose to review this journal because it is a flagship journal of the field with more than 30 years of publications, and BSA and the PDC have been shown to be an integral part of this field; no other publication possessed that form of history with these tools. Lastly, it is possible that the definition of BSA as a multilevel analysis or the use of specific assessment tools was too restrictive.

In general, there remains a need to determine whether preintervention diagnoses such as BSA and the PDC identify more disconnects and whether they lead to better interventions. The fact that such a determination has never occurred is concerning given that members of the field often take great pride in the OBM approach being empirically driven, but the recommended assessment tools of the field lack empirical support. Overall, this is an area rich for confirmation through experimentation and refinement of tools and techniques, assuming that they actually do contribute to substantially improved interventions. The mission of scientific inquiries is to allow one’s respective field to better describe, predict, understand, and control the phenomenon under study. The scientific process, although quite cumbersome and difficult, allows scientists to produce a number of socially important outcomes, namely, empirical descriptions, predictive models, and an understanding of causality. This last outcome, perhaps one of the most important outcomes produced by science, needs to be supported by careful experimentation. The results of this discussion suggest that there is an unfortunate disconnect between the scientific mission and the use of preintervention diagnostic tools. Fortunately, acknowledging and understanding such a disconnect is the first step toward remedying the problem. After all, as the science writer Carl Sagan (1996) stated,

If we were not aware of our own limitations, though, if we were not seeking further data, if we were unwilling to perform controlled experiments, if we did not respect the evidence, we would have very little leverage in our quest for the truth. (p. 263)
ACKNOWLEDGMENT

The authors would like to thank Brittany Barnett and Michael Fantetti for their help with data collection.

REFERENCES


