Are Performance Improvement Professionals Measurably Improving Performance? What *PIJ* and *PIQ* Have to Say About the Current Use of Evaluation and Measurement in the Field of Performance Improvement

Ingrid Guerra-López and Hillary N. Leigh

The ability to prove that performance improvement professionals have made a measurable contribution to their clients and the field remains uncertain (Kaufman & Clark, 1999). Clark and Estes (2000) noted that highly regarded research groups who surveyed performance improvement solutions found "a huge gap between what we think we accomplish and what scientific analyses say we accomplished" (p. 48). Here are some of the findings cited by Clark and Estes (2000) from the work of the National Academic of Sciences and the National Research Council and other independent research groups:

 Scientific studies of training found training interventions often leave participants worse off than before the training intervention (more con-

Measurement and evaluation are at the core of reliably improving performance. It is through these central mechanisms that performance improvement professionals are able to demonstrate the true worth of their efforts. However, the true value of the contributions they make is inconclusive. This article presents a content analysis of 10 years' worth of Performance Improvement and Performance Improvement Quarterly articles as an initial data point to be used for professional reflection and further exploration into the intentions and practices of performance improvement practitioners.

fused, less able to remember important information, less able to use their work-related knowledge effectively).

- More than half of organizational change initiatives are quickly abandoned.
- Kirkpatrick's level one evaluation, the most commonly used method for evaluation, often gives about as much inaccurate information as it does accurate information, including the perception that the object of evaluation has helped, when if fact it has done quite the contrary.
- Studies show that employee empowerment strategies have minimal success in some organizations, and negative consequences in others.



P E R F O R M A N C E I M P R O V E M E N T Q U A R T E R L Y , 2 2 (2) P P . 97 – 1 1 0 © 2009 International Society for Performance Improvement Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/piq.20056 The more rigorous the evaluation, the less likely one is to find evidence of success.

- Myriad studies have found no evidence that multimedia, Internet, and intranet training produce additional learning benefits beyond those already furnished by traditional media such as human trainers or manuals.
- Studies indicate that one-third of the feedback strategies employed in our field do not improve performance, and another third make performance worse.
- Experiments that check for transfer of performance solutions show that even though they work once, they almost never work in other organizational contexts. Because we do not evaluate solutions that may have worked for someone else in another organizational context, we remain ignorant of this failure to transfer.
- Successful performance improvement strategies do exist; however, they are seldom integrated into our most popular performance solutions.

Clark and Estes (2000) also argue that performance improvement professionals tend to "scientize" craft solutions by citing research and evaluation that is often irrelevant or poorly designed. This could suggest a number of things, chiefly that (1) performance improvement professionals do not know how to integrate appropriate research and evaluation practices and findings into their work, or (2) they do not want to integrate appropriate research and evaluation, or (3) they are unaware of the importance of integrating appropriate research and evaluation practices into their work.

This challenge is also faced by other fields closely related to performance improvement. For instance, the fields of applied behavior analysis (ABA) and organizational behavior management (OBM) have "also faced the challenge of extrapolating basic experimental research findings to the behavior of individuals at home, school, work, and in the community" (Culig, Dickinson, McGee, & Austin, 2005, p. 36).

Dickinson (2000) cites various early studies (for example, Andrasik, 1979; Frederiksen & Johnson, 1981; Frederiksen & Lovett, 1980; Hopkins & Sears, 1982; O'Hara, Johnson, & Beehr, 1985) that support the effectiveness of early organizational behavior management interventions. But Dickinson also states that OBM requires increased measurement of social validity, costbenefit analyses, and employee satisfaction or resistance, as well as the long-term effects of interventions. Moreover, the author also cites Balcazar and others (1989), who argue that there is a lack of "largescale interventions in which behavior principles are employed to change the 'cultural foundations' of an organization" (p. 36).

If performance improvement professionals are going to be taken seriously in the scientific and professional communities, they must be able to provide evidence of the rigor and seriousness with which they conduct their work. Publications are artifacts that reflect what a field is about and where its members place their priorities. As such, we want our publications to reflect the true intent, and we hope the actual practice, of the performance improvement field. A review of a couple of key publications in the field might furnish an initial data point that could motivate further inquiry and reflection within the field. In turn, various sources of evidence could be used to produce actionable recommendations that will strengthen the field's contributions and credibility.

Performance Measurement and Evaluation

Predictably improving performance depends not only on setting performance goals, and certainly not only on implementing solutions, but also on continuously tracking progress toward desired goals and taking corrective actions as required. This is why we set goals in the first place, to declare a direction and track our course to ensure we are still heading in the right direction. This is essentially the role of performance evaluation and measurement.

In one study, performance improvement professionals agreed that identifying or verifying performance goals and objectives, at various organizational levels, is the basis for what they do and how they do it, even if they do

not incorporate this into their work as much as they know they should (Guerra, 2003). What appears to be less obvious, or at least not as popular, is that evaluation and measurement are equally important for improving performance. In the same study, participants attributed less importance to evaluation, particularly with regard to evaluating organizational impact on society (customers, local community, environment), as suggested by Kaufman (2006).

Evaluation and measurement certainly do not occur merely at the end of implementing an intervention. Rather, these are integral tools for managing and improving performance at every stage of our work. In a classic work, Rummler and Brache (1995) describe the Predictably improving performance depends not only on setting performance goals, and certainly not only on implementing solutions, but also on continuously tracking progress toward desired goals and taking corrective actions as required.

management function at the organizational level as one that "involves obtaining regular customer feedback, tracking actual performance along the measurement dimensions established in the goals, feeding back performance information to relevant subsystems, taking correction action if performance is off target, and resetting goals so that the organization is continually adapting to external and internal reality" (p. 21). The authors use similar descriptions for the management function at the process level, highlighting the central role of measurement in performance improvement.

The Role of Measurement in Making Decisions. According to *Webster's* dictionary, measurement is estimation of an exact standard. Essentially we use measurement instruments (for example, observation protocols, extant

data review protocols, questionnaires) to compare the object of measurement (a process, an intervention, a project, and so on) to some prespecified standard (such as goals and criteria).

Measurements give us data that we can then turn into intelligence, and in turn we use intelligence to make sound decisions about what to improve and how. Though fundamental to sensible decision making, the most neglected aspect of decision making in the literature is intelligence gathering (Eisenhardt, 1998; Nutt, 2006). Decision making begins when stakeholders see a triggering trend (such as declining revenues or sales) or event (a threat to unionize) as significant, prompting steps to obtain intelligence (Nutt, 2006). In the performance improvement field, we would say that this triggers a needs assessment and analysis, where we can measure gaps in results and establish causal factors that then give us insight as to what solutions are likely to improve performance.

This notion is in fact supported by researchers outside the performance improvement field, who suggest that signals should be decoded as performance gaps (Pounds, 1969; Nutt, 1979; Cowan, 1990), and that the gap will be considered significant if an important performance indicator, such as market share or revenue, falls below a preset criterion; conversely, the signal would be ignored if performance equals or exceeds the expected performance criterion. When a performance gap is detected, it also reveals the magnitude of the concern to be overcome (Cowan, 1990); this magnitude can be a major consideration in prioritizing performance problems for resolution. Decision making is then undertaken to find ways to deal with closing the performance gap, and reduce or eliminate the concern.

The Role of Measurement in Performance Improvement Processes. We typically conduct measurement in the context of needs assessment, causal analysis, evaluation, and research. Performance measurement in the context of needs assessment allows us to determine the gaps between current and desired performance goals, and in the context of *summative evaluation* it enables us to determine whether we have reduced or eliminated these gaps through the performance solutions that were implemented.

Further, overarching processes such as strategic planning and management entail these aforementioned processes and thus logically also entail measurement to a great extent. Finally, if we clarify the purpose of evaluation, we will also appreciate that evaluation can and should occur at every stage of performance improvement (*formative evaluation*); again, by extension, measurement is at the heart of all that we do.

The Value of Evaluation. Although some rightly say that the fundamental purpose of evaluation is determination of the worth or merit of a program or solution (Scriven, 1967), the ultimate purpose, and value, of determining this worth is in making data-driven decisions that lead to improved performance (Guerra-López, 2007). The notion that evaluation's most important purpose *is not to prove but to improve* is an idea originally put forward by Egon Guba decades ago (Stufflebeam & Shinkfield, 2007). Kaufman has similarly

proposed that evaluation data should be used to fix rather than blame (Kaufman & Thomas, 1980). Along these lines, evaluation is simple:

- It compares accomplished results with planned and expected results.
- It can be used to find drivers and barriers to expected performance.
- It should produce actionable recommendations for improving processes, programs, and solutions so that expected performance is achieved or maintained (Guerra-López, 2008).

Not only can ongoing measurement and evaluation help us track, manage, and improve performance, but it is also through performance evaluation and measurement that we are able to prove the worth of our contributions to our clients and to the field. Beyond our ability to sell and implement solutions, the worth of our contributions is evidenced by the measurable results and ultimate impact that we document.

Purpose of the Study. On this view, measurement and evaluation are at the heart of performance improvement. This study was intended as a preliminary step toward understanding their current role within the field of performance improvement. Though there are a variety of indicators that could have been chosen, this study focused on the professional literature as an indication of the importance attributed to evaluation and measurement. For the purposes of this study, the *Performance Improvement* journal (*PIJ*) and *Performance Improvement Quarterly* (*PIQ*), as two of the most representative journals of performance improvement interests and foci in terms of the readership and contributing authors, were used as sources. However, the reader is cautioned to keep in mind that these two journals alone do not fully represent the entire field, or its practitioners. Future studies that look at other relevant journals in performance improvement are warranted to confirm the findings of this study.

The key questions this study sought to answer were:

- To what extent do *PIJ* and *PIQ* reflect an emphasis on performance measurement?
- Is there a difference in the proportion of articles focused on evaluation and performance measurement between the applied journal (*PIJ*) and the research journal (*PIQ*)?
- What is the most prevalent format for these articles?
- Of those articles that emphasize performance measurement, to what extent do they focus on organizational performance as opposed to measurement of specific solutions?

Method

This study used content analysis as a means for answering the research questions. As a research method, content analysis has seen increased sophistication and use in organizational studies (Duriau, Reger, & Pfarrer, 2007). Examples of a similar trend in the field of performance improvement are exhibited in Klein's study (2002) of empirical research in the field and the Marker, Huglin, and Johnsen (2006) replication of it.

According to Duriau and colleagues (2007), content analysis recognizes the relationship between language and attention. This relationship may be examined by various methods, though content analyses generally involve four processes: (1) data collection, (2) coding, (3) analysis of content, and (4) interpretation of results. This study was conducted according to this framework and included an initial phase of content selection.

Content Selection and Data Review. This study reviewed articles from *Per-formance Improvement* and *Performance Improvement Quarterly* published over a 10-year period (1997–2006). In an effort to avoid overreliance on personal opinion, it excluded editorials, interviews, and guest essays from analysis. This resulted in inclusion of 792 articles in the review: 545 from *PIJ*, and 247 from *PIQ*.

Once articles were identified, the data were collected, coded, and analyzed concurrently. These processes occurred in two phases. During the first phase, the article abstracts (or in the case of *PIJ* articles, executive summaries) were reviewed in an effort to answer the research questions. When the review of an abstract yielded an inconclusive response to any of the research questions, the article was set aside for a second phase of more indepth data review of the entire article. Of these, only three articles were deemed difficult to classify and set aside for a final and third phase of review approximately a week later as a check for consistency. All of the data review was performed by the same individual; however, another individual was consulted during the second and third phases of review in order to come to a consensus about final coding. All phases of data analysis were conducted using operational definitions that were carefully discussed and defined by both researchers.

Attention was defined at two levels: emphasis and focus. In this case, performance measurement emphasis was indicated by devotion to the topic of measurement in the article's (1) abstract (or executive summary); (2) statement of intent; or (3) titles of the headings, tables, figures, or charts included within it. Most of this study was concerned with the emphasis-level of attention. The issue of *focus* was relevant only for research question four (the extent to which authors addressed organizational performance as opposed to measurement of specific interventions). Focus was judged on the basis of which level of evaluation the article advocated; when an article discussed measurement both in support of a particular intervention and organizational performance, this distinction was made according to which area had significantly more content dedicated to it. For example, an article that provided step-by-step guidelines on how to evaluate an electronic performance support system (EPSS) and offered general comments of a few sentences about the importance of measuring organizational performance would be considered to have a measurement emphasis and a focus on intervention measurement.

Performance measurement was defined as measurement activities relating to the needs assessment, analysis, strategic planning, evaluation, or performance tracking contexts. More particularly, *needs assessment* is a methodology associated with measuring gaps in results; *analysis* is a methodology associated with breaking down elements of performance to identify causal factors. Strategic planning includes establishing and tracking strategic measures for the purposes of long-term planning, management, and evaluation. *Evaluation* includes determining the effectiveness of organizations as well as specific activities, interventions, or anything else meant to contribute to performance; *performance tracking* involves specific measurement and tracking techniques such as development and use of indicators and scorecards.

Format was concerned with the general nature of an article. Those articles that were *model-oriented* described a particular model for evaluation or measurement; *persuasive* articles advocated use of evaluation or measurement in general, or more specifically the benefits of using a particular model over another. *Methodological* articles illustrated how to apply evaluation or measurement methodology (for instance, case studies or job aids). When an article exhibited the characteristics of more than one format, this distinction was made according to which area had significantly more content dedicated to it.

Findings

For the sake of clarity, findings have been arranged by each research question as shown here.

1. To what extent do *PIJ* and *PIQ* reflect an emphasis on performance measurement?

This research question was organized into several subquestions:

Overall emphasis: PIJ published 188 of 545 articles (34%) related to some aspect of performance measurement, inclusive of the evaluation articles. *PIQ* published 186 of 247 articles (75%) related to some aspect of performance measurement, inclusive of evaluation articles (see Figure 1).

Needs assessment emphasis: PIJ published 15 (2.75%) of 545 articles on the topic of needs assessment, while *PIQ* published 4 (1.62%) of 247.

Analysis emphasis: PIJ published 57 (10.46%) of 545 articles on causal analysis, while *PIQ* published 59 (23.89%) of 247.

Strategic planning emphasis: PIJ published 17 (3.12%) of 545 articles on strategic planning, while *PIQ* published 7 (2.83%) of 247.

Evaluation emphasis: PIJ had 57 evaluation articles out of a total of 545 articles published, so roughly 10% of publications dealt with evaluation. Meanwhile, *PIQ* published 108 evaluation articles out of 247 (44%).

Performance tracking emphasis: PIJ published 42 (7.71%) of 545 articles on performance tracking, while *PIQ* published 8 (3.24%) of 247 (see Figure 2).

2. Is there a difference in the proportion of articles focused on evaluation and performance measurement between the applied journal (*PIJ*) and the research journal (*PIQ*)?

There is an obvious difference, with *PIQ* publishing four times as many evaluation articles as *PIJ*. In terms of performance measurement in general, *PIQ* has published more than twice as many articles as *PIJ*. During the content analysis, the changes in editors were also examined as one possible factor for variations in focus on performance measurement (see Figure 3).

3. What is the most prevalent format for these articles?

Of the 374 articles emphasizing performance measurement, 246 (66%) of them were methodological (they offered guidelines for how to measure a performance, typically through explanation of a job aid or a case study); the remaining 128 articles were evenly divided between presenting a model (n = 64) and advocating evaluation in general or a model in particular (n = 64) (see Figure 4).



FIGURE 1. OVERALL EMPHASIS ON PERFORMANCE MEASUREMENT



FIGURE 2. EMPHASIS ON SPECIFIC PERFORMANCE MEASUREMENT CONTEXTS







FIGURE 4. MOST PREVALENT FORMATS FOR ARTICLES RELATED TO PERFORMANCE MEASUREMENT

4. To what extent are performance measurement articles focused on organizational performance measurement as opposed to measurement of a specific solution?

In *PIJ*, 90 of 188 articles (48%) focused on measurable performance indicators. In *PIQ*, 47 of 186 (25%) were focused on measurable performance indicators rather than theoretical, more abstract variables.

In *PIJ*, 90 (48%) of 188 focus on performance measurement, while 98 (52%) focus on the actual solution. *PIQ* published 47 (25%) of 186 that focused on performance measurement, and 139 (75%) related to measurement of a specific solution (see Figure 5).

Discussion

Considering that formative evaluation can play an important role in every performance management activity, finding only one-tenth of *PIJ* publications with an evaluation component seems harshly disproportionate.



FIGURE 5. TO WHAT EXTENT DO PERFORMANCE MEASUREMENT ARTICLES FOCUS ON ORGANIZATIONAL VERSUS SPECIFIC SOLUTIONS?

One would not necessarily expect evaluation to be the central focus of every *PIJ* article; however, one might expect to see some evaluative component in much more than one tenth.

PIQ significantly surpassed *PIJ* in this area, with close to half its articles reflecting an evaluation focus. Indeed, evaluation research would naturally be thought to belong in this research journal, yet not all evaluations fall into this category. As the practitioner publication, one could have still expected to see more evaluation articles in *PIJ* that were perhaps less formal or structured.

The performance measurement article ratios are expectedly higher, because for the purposes of this study we considered evaluation as a particular category within performance measurement. One-third of PII articles were indeed directed at some aspect of performance measurement, suggesting relative awareness of measurement as an important component of performance management and improvement. The PIQ results illustrated an outstanding two-thirds of articles focused on measurement. This certainly does not come as a surprise; measurement is a central mechanism in research, and PIQ is a key research journal of the field. If we follow this argument, however, it is then interesting that the findings did not reveal an even higher ratio. One issue that could account for not finding more PIQ articles focused on measurement is that the remaining articles were for the most part qualitative case studies. Both qualitative and quantitative studies have the potential to be equally strong and valid, provided they are conducted rigorously and appropriately. However, even though quantitative research is generally centered on measurement, qualitative research is generally centered on description.

Because *PIJ* tends to be considered the practitioner journal in the performance improvement field while *PIQ* is considered the research

journal, it may not be too surprising to find significantly higher "talk" of measurement in the research journal. Yet measurement does not exclusively take place in a traditional research context. Measurement is at the core of what all serious performance consultants do (Rummler, 2004). If this is true, then why are their evaluation and measurement practices not captured in their articles, even at an anecdotal level? We should at least expect to see publications about measurement practices as much as we see publications on interventions. After all, one of the characteristics that set us apart from other types of consultants (and indeed from typical salespeople) should be the emphasis on actual improvement ... not just selling or trends in implementation.

The primary differences between *PIQ* and *PIJ* should not necessarily be the topics themselves but rather whether the topics are being discussed in the context of a research study or in a more applied setting. Evaluation and measurement can and should be done in both contexts.

The fact that almost two-thirds of measurement articles are focused on giving the readership how-tos is a promising indication that there are indeed performance improvement professionals advocating and disseminating measurement practices. We cannot assume that these are best practices, but it is certainly a step in the right direction and could be used as the first awareness-raising step in encouraging the readership to follow up on how to integrate measurement into their practices.

Essentially, *PIJ* has a half-and-half split between articles focused on measuring organizational performance and those measuring the performance of a specific solution. The focus on measurement in *PIQ*, by contrast, is primarily focused on testing the performance or effectiveness of a specific solution. This may not be an unexpected result; *PIQ* publishes research articles, and by their very nature they require articulation of specific variables. Most of the time, these independent variables—that is, the interventions—are tested for specific impact, rather than focusing on general organizational performance variables.

In regard to the effect of editorial change on performance measurement emphasis, two of three decreases in *PIJ* occur in years where there was a change in editor. By contrast, *PIQ* has fewer editorial changes. Furthermore, given *PIQ's* peer review process for selecting articles, one would suspect the journal's emphasis is less susceptible to editor change. Additionally, if we consider special issues and guest editing, editor change does not paint the whole picture.

Conclusion and Recommendations

This study was intended as a preliminary step in uncovering the perceived value of performance evaluation and measurement for performance improvement professionals. The data, in particular those related to our practitioner journal, *PIJ*, reveal that our attention to evaluation and measurement is not at a level that supports our claims to add measurable value to our clients. Its publications show relatively low attention to performance measurement, and even less attention to evaluation in particular. If evaluation and measurement are not central mechanisms of our performance improvement projects, how are we supporting our claims to add measurable value to our clients?

Performance data tracked through evaluation and measurement should be readily available at a given time for any of our projects, even if "the client will not pay for it." This study also included a preliminary consideration of the impact of editor change on the measurement and evaluation focus within the journals. It is likely that editors' opinions and values influence the selection of articles, but these data do not conclusively support a causal linkage between the two. Rather, the data suggest that more study in this area is warranted, especially as it may guide editorial selection practices.

It would behoove us performance improvement professionals to not only talk about evaluation and measurement but actually integrate it into our work. If we want future analysis of our practices and solutions to support our professional and ethical credibility, we must have the data to support our claims. Performance data tracked through evaluation and measurement should be readily available at a given time for any of our projects, even if "the client will not pay for it." Tracking the effectiveness of our own work should be ingrained into what we do anyway. For example, in the context of managing our projects, we must track what was planned against what was accomplished. It does not necessarily have to be considered a separate activity and deliverable for the client, though the client certainly will benefit in many ways, including justifying decisions with relevant, reliable, valid, and not least of all well-documented data. No matter what we call it, "evaluation" or "performance tracking" can be integrated as an important part of all phases of our projects, including the final products and contributions we deliver.

It is worth reiterating that the literature is but one indicator of the practices of a field. As with interpretation of any other dataset, it must be verified and understood in the context of other relevant, reliable, and valid indicators before confidently making interpretations and drawing conclusions (Guerra-Lopez, 2008).

Readers are also cautioned to keep in mind that whereas operational definitions and a review protocol were used to categorize and count the articles, the findings are dependent on subjective judgments; they could have been different had another observer used the same tools to make the observations. Replication of this study would be advisable in order to confirm findings.

Finally, it must be noted that *PIJ* and *PIQ* do not necessarily represent the values, intentions, and practices of every performance improvement professional. Rather, the articles reviewed for this study are more directly a reflection of the authors, editors, and editing review boards that accept manuscripts for publication.

References

- Andrasik, F. (1979). Organizational behavior modification in business settings: A methodological and content review. *Journal of Organizational Behavior Management*, 2, 85–102.
- Balcazar, F. E., Shupert, M. K., Daniels, A. C., Mawhinney, T. C., & Hopkins, B. L. (1989). An objective review and analysis of ten years of publication in the *Journal of Organizational Behavior Management. Journal of Organizational Behavior Management*, 2, 7–37.
- Clark, R., & Estes, F. (2000). A proposal for the collaborative development of authentic performance technology. *Performance Improvement*, *39*(4), 48–53.
- Cowan, D. A. (1990). Developing a classification structure of organizational problems: An empirical investigation. *Academy of Management Journal*, 33(2), 366–390.
- Culig, K. M., Dickinson, A. M., McGee, H. M., & Austin, J. (2005). An objective comparison of applied behavior analysis and organizational behavior management research. *Journal of Organizational Behavior Management*, 25(1), 35–72.
- Dickinson, A. M. (2000). The historical roots of Organizational Behavior Management in the private sector: The 1950s–1980s. *Journal of Organizational Behavior Management*, 20(3/4), 9–58.
- Duriau, V. J., Reger, R. K., & Pfarrer, M. D. (2007). A content analysis of the content analysis literature in organizational studies: Research themes, data sources, and methodological refinements. Organizational Research Methods, 10(1), 5–34.
- Eisenhardt, K. (1998). Decision making and all that jazz. In V. Papadakis & P. Barwise (Eds.), Strategic decisions. Boston: Kluwer.
- Frederiksen, L. W., & Johnson, R. P. (1981). Organizational behavior management. *Progress in Behavior Modification*, 12, 67–118.
- Frederiksen, L. W., & Lovett, S. B. (1980). Inside organizational behavior management: Perspectives on an emerging field. *Journal of Organizational Behavior Management*, 2, 193–203.
- Guerra, I. (2003). Key competencies required of performance improvement professionals. *Performance Improvement Quarterly*, 16(1).
- Guerra-López, I. (2007). Evaluating impact: Evaluation and continual improvement for performance improvement practitioners. Amherst, MA: HRD Press.
- Guerra-López, I. (2008). Performance evaluation: Proven approaches for improving program and organizational performance. San Francisco: Jossey-Bass.
- Hopkins, B. L., & Sears, J. (1982). Managing behavior for productivity. In L. W. Frederiksen (Ed.), Handbook of organizational behavior management (pp. 393–425). New York: John Wiley.
- Kaufman, R. (2006). Change, choices, and consequences: A guide to Mega thinking and planning. Amherst, MA: HRD Press.
- Kaufman, R., & Clark, R. (1999). Re-establishing performance improvement as a legitimate area of inquiry, activity, and contribution: Rules of the road. *Performance Improvement*, 38(9), 13–18.
- Kaufman, R., & Thomas, S. (1980). Evaluation without fear. New York: New Points/Franklin Watts.
- Klein, J. D. (2002). Empirical research on performance improvement. Performance Improvement Quarterly, 15(1), 99–110.
- Marker, A., Huglin, A., & Johnsen, L. (2006). Empirical research on performance improvement: An update. *Performance Improvement Quarterly*, 19(4), 7–22.
- Nutt, P. (1979). Calling out and calling off the dogs: Managerial diagnoses in organizations. Academy of Management Review, 4(2), 203–214.
- Nutt, P. (2006). Intelligence gathering for decision making. Omega, No. 25, 604–622.
- O'Hara, K., Johnson, C. M., & Beehr, T. A. (1985). Organizational behavior management in the private sector: A review of empirical research and recommendations for further investigation. *Academy of Management Review, 10*, 848–864.
- Pounds, W. (1969). The process of problem finding. *Industrial Management Review*, 1969, 1–19.

- Rummler, G. A. (2004). Serious performance consulting: According to Rummler. Silver Spring, MD: International Society for Performance Improvement.
- Rummler, G. A., & Brache, A. P. (1995). *Improving performance: How to manage the white space on the organization chart* (2nd ed.). Jossey-Bass: San Francisco.
- Scriven, M. (1967). The methodology of evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.), *Perspectives on curriculum evaluation*. New York: McGraw-Hill.
- Stufflebeam, D., & Shinkfield, A. (2007). *Evaluation theory, models, and applications*. San Francisco: Wiley.

INGRID GUERRA-LÓPEZ

Ingrid Guerra-López is an associate professor at Wayne State University, associate research professor at the Sonora Institute of Technology in Mexico, and director of the Institute for Learning and Performance Improvement. Her research, teaching, and consultancy focuses on improving performance and management decision making through design, development, and use of performance measurement, tracking, and management systems. She is also principal of Intelligence Gathering Systems. She has written five books on performance evaluation and assessment, as well as publishing nearly 30 articles and 10 book chapters on performance improvement, assessment, and evaluation. *E-mail*: iguerra@wayne.edu

HILLARY N. LEIGH

Hillary N. Leigh is a doctoral student in Wayne State University's instructional technology program. Her research interests include evidence-based intervention selection and justification for the field of performance improvement. Her dissertation topic relates to practitioners' usage of scientific and artistic evidential sources in selecting an intervention. She has consulted with health care, educational, and retail organizations to select, develop, implement, and evaluate a variety of instructional and noninstructional interventions. *E-mail:* hleigh@wayne.edu