Empirical Research on Performance Improvement

James D. Klein
Arizona State University

ABSTRACT

Renewed interest in research on performance improvement provides an ideal opportunity to examine how much empirical work is actually published and to identify the focus of research in the field. This paper describes a study conducted to examine the empirical research on performance improvement. The study involved a content analysis of articles published in *PIQ* from 1997 through 2000. Results indicated that empirical research accounted for about one-third of all articles published in *PIQ*. A detailed analysis of the empirical research studies published in the journal is provided.

The importance of empirical research to performance improvement has been a focus of discussion among leaders in the human performance technology (HPT) field in recent years. Some individuals indicate that research and practice must be integrated if findings are to add value to HPT (Brethower, 2000). Others suggest that the empirical foundations of the field have not kept pace with HPT practices and call for increased, targeted research activity (Stolovitch, 2000; Sugrue & Stolovitch, 2000). Still others believe that HPT is in danger of becoming a "craft" because many professionals identify solutions without collecting empirically based research data (Kaufman & Clark, 1999).

As a field of practice, HPT relies on empirical research to improve human performance. HPT achieves desired performance through approaches that have been derived from scientific research or documented evidence (Stolovitch &

Keeps, 1999). Practitioners conduct research to identify performance problems, their causes and solutions to obtain formative data while an intervention is being developed, and to determine if an intervention is successful. Furthermore, academic researchers often collect data to synthesize trends and discover new models, methods, and technologies for performance improvement (Foshay, Moller, Schwen, Kalman, & Haney, 1999).

Practitioners and academic researchers use a variety of alternative paradigms, methods, and tools to conduct research on performance improvement. Dean (1999, p. 325) indicated, "The HPT professional uses both quantitative and qualitative data, as needed…uses observation, not hearsay to collect facts…[and] relies on direct, comparative, and economic measures." Foshay et al. (1999) recommended that investigation of HPT questions require the careful selection and combination of

research methods. Binder (1995) advised that HPT researchers should use a variety of quality, quantity, and cost measures.

Renewed interest in research on HPT provides an opportunity to examine how much empirical research is actually being published in the performance improvement literature and to determine the focus of this research. A recent review of articles from four professional publications that focus on HPT and training (*Performance Improvement, Performance Improvement Quarterly, Training and Development,* and *Technical Training Magazine*) revealed that empirical research accounted for only 7% of all articles published (Werner & Klein, 2000). Furthermore, two recent reviews of *Educational Technology Research and Development (ETR&D)* have shown that empirical research accounted for less than a third of the articles published in the instructional design and development section of the journal (Driscoll & Dick, 1999; Klein, 1997).

The purpose of this article is to describe the results of a study conducted to determine the focus of empirical research on performance improvement. The study involved an examination of articles published in *Performance Improvement Quarterly* (*PIQ*)[1], the scholarly journal of the International Society for Performance Improvement (ISPI). [Footnotes are located at the end of the article.] According to Dean (see Sugrue & Stolovitch, 2000), *PIQ* extends HPT as a field of study through the publication of original scholarly articles including experimental studies, literature reviews, theoretical concepts, and case study applications.

The current study was designed to answer the following questions:

1) How much empirical research is being published in *PIQ*?
2) What is the focus of this research?
3) Which research methods are being employed to study performance improvement?
4) Which performance improvement interventions are being empirically tested?
5) How are these interventions being evaluated?

> *The importance of empirical research to performance improvement has been a focus of discussion among leaders in the human performance technology field in recent years.*

Method

To answer these questions, a content analysis of articles published in *PIQ* from 1997 through 2000 was conducted. A total of 138 articles from 15 issues of the journal were analyzed. Articles published in Volume 10, Number 1 (1997) were excluded from the content analysis since that issue of the journal reprinted articles previously published in *PIQ*. Book reviews, bibliographies, and abstracts were also excluded from the analysis.

First, each article was classified into one of four categories based on research conducted by Dick and Dick (1989), Higgins, Sullivan, Harper-Marinick, and Lopez (1989), and Klein (1997). These categories were:

- *Description*–provides information about a specific project, method, course, or intervention with no use of data to draw conclusions
- *Literature review*–summarizes a body of literature as a critique or to draw implications for practice
- *Empirical Research*–reports on a research study that used data to draw conclusions
- *Editorial/Commentary*–states the opinions of an editor, guest editor, or other individual toward a particular topic, trend or issue.[2]

Next, each article classified as an empirical research study was examined to identify its topic focus. Empirical research studies were also analyzed using a classification scheme developed by Driscoll and Dick (1999) to determine which research method was employed. The research methods in this classification scheme were:

- *Case Study*–a study of how or why something occurs
- *Evaluation*–a study to determine the impact of a program or intervention
- *Experiment*–a study of the effect of manipulated treatment variables on observed variables
- *Qualitative-Naturalistic*–a study aimed at developing an understanding of a human system, involving the collection of non-numerical data and rich descriptions of natural events

- *Survey*–a study describing the distribution of responses to a questionnaire

Empirical research articles were further analyzed to determine whether a performance improvement intervention was implemented and empirically tested. Interventions were classified by type and examined to determine how they were evaluated (see Table 1). Interventions were analyzed using Kirkpatrick's (1994) four-level model of evaluation expanded to include societal consequences (Kaufman & Keller, 1994; Watkins, Leigh, Foshay, & Kaufman, 1998).

Results

Information on the types of articles published in *PIQ* from 1997-2000 is provided in Table 2. These data indicate that empirical research studies accounted for 36% of all articles published in *PIQ*. Articles about a specific project, method, course, or intervention with no data to draw conclusions (i.e., descriptions) made up 31% the articles published. Literature reviews accounted for 18%, and editorials/commentaries made up 15% of the 138 articles analyzed in this study.

Table 3 provides information about the topic focus of the of 49 empirical research studies published in *PIQ* (see the Appendix for a reference list of these studies). These data show that the professional practices of performance technologists and instructional designers was the most frequently researched topic (*n*=9), followed by the topics of strategies for training and instruction (*n*=7), transfer of training and learning (*n*=6), and workplace diversity (*n*=5).

Table 1
Evaluation Levels and Tools

Level	Description	Tools
Reaction	Attitudes toward the intervention	Surveys, interviews, focus groups
Learning	Attainment of skills, knowledge & competence	Performance and achievement tests
Behavior	Job performance, application, transfer	Direct observation, surveys, interviews
Results	Impact of intervention on the organization	Return-on-investment, cost-benefit analysis
Societal Benefit	Impact of intervention on society	Cost-consequences analysis

Table 2
Types of Articles Published in *PIQ* from 1997–2000*

Article type	Number	Percentage
Editorial/Commentary	21	15%
Literature Reviews	25	18%
Descriptions	43	31%
Empirical Research	49	36%

*A total of 138 articles were analyzed

A list of the research methods employed in the 49 empirical studies published in *PIQ* can be found in Table 4. Survey research (31%) and case study methods (22%) were most frequently used in the empirical studies published in the journal. Experimental research accounted for 18%, and evaluation studies made up 16% of the empirical articles.

Results of the content analysis also revealed that 24 out of the 49 research studies published in *PIQ* focused on the implementation and empirical test of a performance intervention. (The Appendix identifies these articles). Twenty-three of these studies examined a training or instructional intervention (see Table 5). Most of these implemented classroom instruction (n = 17). Only one study investigated a non-instructional solution. In addition, more than half (n=14) examined behavior

Table 3
Topics of Research Articles Published in *PIQ* from 1997–2000*

Article topic	Number
Practices of HPT/ID professionals	9
Training and instructional strategies	7
Transfer of training/learning	6
Workplace Diversity	5
Action Learning	3
Incentives/Reinforcement/Feedback	3
Performance/Needs Analysis	3
Return-on-Investment	3
Organizational Culture	3
Management & Leadership	2
Coaching/Facilitation	2
Total Quality Management	1
Distance Learning	1
Employee Selection	1

*49 research articles were analyzed

such as on-the-job performance and transfer of training. The analysis also revealed that:

- Surveys and questionnaires were used most often to measure reaction.
- Performance assessments were used most often to measure learning.
- Interviews and surveys were used most often to measure perception toward application, transfer, and on-the-job performance.
- Direct observation was seldom used to measure the impact of training to on-the-job performance.
- Actual work performance was used to evaluate the impact of a non-instructional intervention (incentive system).
- Cost-benefit analysis was used to measure the impact of on-the-job-training.
- Pre- to post- measures such as profit margin, sales, and revenue were obtained to evaluate the re-turn-on-investment of one training program.
- Cost-consequence analysis was used to determine the impact of a workforce development program on society.

Table 4
Research Methods Employed in Studies Published in *PIQ**

Type of inquiry	Number	Percentage
Survey	15	31%
Case study	11	22%
Experimental	9	18%
Evaluation	8	16%
Naturalistic	3	6%
Content analysis	3	6%

*49 empirical studies were analyzed

Discussion

The field of human performance technology is grounded in empirical research (Stolovitch & Keeps, 1999). Recently, leaders in the field have called for more research on HPT (Brethower, 2000; Kaufman & Clark, 1999; Stolovitch, 2000; Sugrue & Stolovitch, 2000. Unfortunately, the current study suggests that appeals for empirical research are going unheeded by many authors.

Results indicated that articles using data to draw conclusions (i.e., empirical research) accounted for a little more than one-third of the papers published in *PIQ* from 1997 to 2000. When compared with the results of two recent content analyses of the instructional design and technology (IDT) literature (Driscoll & Dick, 1999; Klein 1997), the current findings show that *PIQ* is publishing empirical research at approximately the same rate as other scholarly journals in the field. However, when the empirical basis of HPT and IDT is considered, these findings might be interpreted as less sanguine.

Additional empirical work on a variety of HPT topics is undoubtedly appropriate. Yet the current study suggests that more data-based research on topics such as employee selection, distance education, coaching, facilitation, and feedback is essential. It is interesting to note that during the four-year time period covered by this analysis, *PIQ* devoted two complete issues on action learning and another on distance education. However, very little empirical work was published on those two topics, especially when one considers the number of journal pages devoted to them.

Rigorous empirical research on all types of performance interventions is also required to validate their effects on individuals and organizations (Farrington & Clark, 2000; Kaufman & Clark, 1999). Less than half of the research studies published in *PIQ* focused on the implementation and empirical test of a performance intervention; all of these studies except one investigated a training or instructional intervention. Recently, Brethower (1999, p. 319) noted:

The fact is that many HP technologists view training as a last resort, to be employed only when no other means of achieving improved performance will work. Even when instruction is required, it is frequently only one among a number of interventions employed to address a problem or realize an opportunity.

The current study suggests that more research on the effects of noninstructional performance interventions should be conducted and published in the literature. Rigorous studies on interventions such as performance support systems, job aids, knowledge management, motivation, and incentive systems is required to inform the field.

According to Foshay et al. (1999), the effect of any performance intervention should be judged by cumulative changes in individual behavior. Findings from the present study revealed that in the fourteen cases where on-the-job performance and transfer of training were examined, researchers relied mostly on self-report measures of these outcomes rather than on direct observation. While observation is more costly and time-consuming than surveys and interviews, the increased use of direct measures of job performance and transfer would help to inform practitioners about the actual benefits of a particular intervention.

Furthermore, research studies should also examine whether a performance intervention adds value to organizations and to society. The findings of this review of *PIQ* and a recent review of *Performance Improvement* (Guerra, 2001) indicate

Table 5
Types of Performance Interventions and Levels of Evaluation*

Intervention Type		Reaction	Learning	Behavior	Results	Societal Impact
Classroom instruction	(*n*=17)	6	9	12	1	0
Web-based instruction	(*n*=3)	2	3	1	0	0
Self-paced workbooks	(*n*=1)	0	1	0	0	0
On-the-Job Training	(*n*=1)	0	0	0	1	0
Workforce Development	(*n*=1)	0	0	0	0	1
Incentive System	(*n*=1)	0	0	1	0	0

Evaluation Level

*24 articles reported on the implementation of a performance intervention; 12 focused on more than one level of evaluation.

that very few published articles include data on cost benefit or return-on-investment.

Finally, additional studies should be conducted to examine the focus of empirical research on performance improvement. The current study analyzed a sample of published research in the field. Future research should investigate the actual work practices of performance technologists to determine which interventions are being empirically tested and which research methods and tools are being employed to study performance improvement. Continued interest in the empirical research on performance improvement will help validate and refine the field.

[1] While an examination of *PIQ* provides a representative view of scholarly articles on performance improvement, I recognize that it does not necessarily provide a comprehensive one. Other journals such as *HRD Quarterly* publish scholarly research on HPT. I chose to examine *PIQ* because of its primary focus on performance improvement.

[2] This category was not included in the studies by Dick and Dick (1989), Higgins et al. (1989), and Klein (1997). I included it in the current study because editorials/commentaries that include scholarly literature are regularly published in *PIQ*.

Note: The author wishes to acknowledge Jayne Klein and Jenny Lynn Werner for their assistance with this study.

References

Binder, C. (1995). Promoting HPT innovation: A return to our natural science roots. *Performance Improvement Quarterly , 8*(2), 95-113.

Brethower, D. (1999). Human performance interventions of a noninstructional nature. In H. Stolovitch & E. Keeps (Eds.), *Handbook of Human Performance Technology* (pp. 319-320). San Francisco: Jossey-Bass/Pfeiffer.

Brethower, D. (2000). Integrating theory, research and practice in human performance technology. *Performance Improvement, 39*(4), 33-43.

Dean, P. (1999). Designing better organizations with human performance technology and organizational development. In H. Stolovitch & E. Keeps (Eds.), *Handbook of Human Performance Technology* (pp. 321-333). San Francisco: Jossey-Bass/Pfeiffer.

Dick, W., & Dick, D. (1989). Analytical and empirical comparisons of the *Journal of Instructional Development and Educational Communications and Technology Journal*, *Educational Technology Research and Development*, *37*(1), 7-18.

Driscoll, M., & Dick, W. (1999). New research paradigms in instructional technology: An inquiry. *Educational Technology Research and Development, 47*(5), 7-18.

Farrington, J., & Clark, R. (2000). Snake oil, science, and performance products. *Performance Improvement, 39*(10), 5-10.

Foshay, W., Moller, L., Schwen, T., Kalman, H., & Haney, D. (1999). Research in human performance technology. In H. Stolovitch & E. Keeps (Eds.), *Handbook of Human Performance Technology* (pp. 895-914). San Francisco: Jossey-Bass/Pfeiffer.

Guerra, I. (2001). Performance improvement based on results: Is our field adding value? *Performance Improvement, 40*(1), 6-10.

Higgins, N., Sullivan, H., Harper-Marinick, M., & Lopez, C. (1989). Per-

spectives on educational technology, research, and development. *Educational Technology Research and Development*, *37*(1), 81-87.

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., & Keller, J. (1994). Levels of evaluation: Beyond Kirkpatrick. *Human Resources Quarterly*, *5*(4), 371-380.

Kirkpatrick, D. (1994). *Evaluating training programs: The four levels*. San Francisco: Berrett-Koehler.

Klein, J. (1997). ETR&D–Development: An analysis of content and survey of future direction. *Educational Technology Research and Development*, *45*(3), 57-62.

Stolovitch, H. (2000). Human performance technology: Research and theory to practice. *Performance Improvement, 39*(4), 7-16.

Stolovitch, H., & Keeps, E. (1999), What is human performance technology? In H. Stolovitch & E. Keeps (Eds.), *Handbook of Human Performance Technology* (pp. 3-23). San Francisco: Jossey-Bass/Pfeiffer.

Sugrue, B., & Stolovitch, H. (2000). Report of 1999 ISPI symposium: Appropriate inquiry in human performance technology. *Performance Improvement, 39*(1), 33-36.

Watkins, R., Leigh, D., Foshay, R., & Kaufman, R. (1998). Kirkpatrick plus: Evaluation with a community focus. *Educational Technology Research and Development, 46(4)*, 90-96.

Werner, J., & Klein, J. (2000, October). *Performance data in professional publications: Signs on the road less traveled*. Paper presented at the annual meeting of the Association for Educational Communications and Technology, Denver, CO.

JAMES D. KLEIN is professor of Educational Technology at Arizona State University in Tempe and the Development Editor of *Educational Technology Research and Development (ETR&D)*. He is currently a member of the International Board of Standards for Training, Performance, and Instruction (IBSTPI). His teaching, research, and consulting activities are in the areas of instructional systems design and strategies for active learning. *Mailing address*: Division of Psychology in Education, Arizona State University, Box 870611, Tempe, AZ 85287-0611. *Telephone*: 480-965-0349. *E-mail*: james.klein@asu.edu

Appendix

The following 49 articles published in *Performance Improvement Quarterly* from 1997-2000 were classified as empirical research. The 24 articles marked with an asterisk (*) focused on the implementation and empirical test of a performance intervention.

*Bucklin, B., Dickinson, A., & Brethower, D. (2000). A comparison of the effects of fluency training and accuracy training on application and retention. *13*(3), 140-163.

Carliner, S. (1998). How designers make decisions: A descriptive model of instructional design for informal learning in museums. *11*(2), 72-92.

*Cruz, B. (1997). Measuring the transfer of training. *10*(2), 83-97.

*Dabbagh, N., Jonassen, D., Yueh, H., & Samouilova, M. (2000). Assessing a problem-based learning approach to an introductory instructional design course: A case study. *13*(3), 60-83.

Ellinger, A. (1999). Antecedents and consequences of coaching behavior. *12*(4), 45-70.

Ellinger, A., Keller, S., & Ellinger, A. (2000). Developing interdepartmental integration: An evaluation of three strategic approaches for performance improvement. *13*(3), 41-59.

*Foxon, M. (1997). The influence of motivation to transfer, action planning, and manager support on the transfer process. *10*(2), 42-63.

Goins, S., & Mannix, E. (1999). Self-selection and its impact on team diversity and performance. *12*(1), 127-147.

*Graham, S., Wedman, J., Tanner, T., & Monahan, C. (1998). Yes, classroom sales training can enhance performance. *11*(2), 101-112.

Hatcher, T., & Ward, S. (1997). Framing: A method to improve performance analyses. *10*(3), 84-103.

*Hemmes, K., Long, C., & Rowland, G. (1998). Situating learning of human performance technology. *11*(3), 16-31.

Hubiak, W. (1997). Mental models of quality. *10*(4), 31-50.

*Jacobs, R., & Hruby-Moore, M. (1998). Learning from failure: A cost-benefit analysis study which resulted in unfavorable financial outcomes. *11*(2), 93-100.

*Jones, N., & Laffey, J. (2000). The diffusion of collaborative technologies into a college classroom. *13*(4), 29-46.

*Julian, M., Kinzie, M., & Larsen, V. (2000). Compelling case experiences: Performance, practice, and application for emerging instructional designers. *13*(3), 164-201.

*Kaufman, R., Watkins, R., Sims, L., Crispo, N., Hall, J., & Sprague, D. (1997). Cost-consequence analysis: A case study. *10*(3), 7-21.

Kaufman, R., Watkins, R., Stith, M., & Triner, D. (1998). The changing corporate mind: Organizations, vision, missions, purposes, and indicators on the move toward societal payoffs. *11*(3), 32-44.

Keith, J., & Gresso, D. (1997). The relationship among leadership behaviors of leaders in training organizations, training methods, and organization profitability. *10*(3), 56-66.

Kolb, J., & Rothwell, W. (2000). Challenges and problems reported by small group facilitators. *13*(4), 122-136.

Korth, S. (1997). Planning HRD interventions: What, why, and how. *10*(4), 51-71.

Korth, S. (2000). Creativity and the design process. *13*(1), 30-45.

Kuchinke, K. (2000). Information and feedback seeking in U.S. and British human resources development and training settings. *13*(1) 46-59.

Kunneman, D., & Sleezer, C. (2000). Using performance analysis for training in an organization implementing ISO-9000 manufacturing practices: A case study. *13*(4), 47-66.

Layng, J. (2000). Diversity at work: A case study of using videotape training to enhance performance. *13*(4), 67-86.

*Lee, C., & Kahnweiler, W. (2000). The effect of a mastery learning technique on the performance of a transfer of training task. *13*(3), 125-139.

*Lee, K., & Pucel, D. (1998). The perceived impacts of supervisor reinforcement and learning objective importance on transfer of training. *11*(4), 51-61.

*Lohman, M. (1997). Effects of an inductive versus a deductive instructional approach on the constructive feedback and problem-solving skills of supervisors. *10*(3), 37-55.

Loughner, P., & Moller, L. (1998). The use of task analysis procedures by instructional designers. *11*(3), 79-101.

*Machin, M., & Fogarty, G. (1997). The effects of self-efficacy, motivation to transfer, and situational constraints on transfer intentions and transfer of training. *10*(2), 98-115.

Marshall, J., & Rossett, A. (2000). An exploratory study of the relationship between knowledge management and performance professionals. *13*(3), 23-40.

Marshall, R., Ungson, G., & Pan, Y. (2000). Organizational learning: A perspective from the choice and pattern of sequential modes of foreign market entry. *13*(2), 117-137.

*Pershing, J., & Lee, S. (2000). Transfer evaluation of knowledge and skill development: A case study of an instructional systems development training program. *13*(1), 7-29.

Richard, O., & Johnson, N. (1999). Making the connection between formal human resource diversity practices and organizational effectiveness: Behind management fashion. *12*(1), 77-96.

Rossett, A., & Tobias, C. (1999). A study of the journey from training to performance. *12*(3), 31-43.

*Schor, S., Sabiers, M., Hall, J., & Anakwe, U. (2000). A management skills course: Did it really make a difference? *13*(3), 5-22.

Sleezer, C., Gregson, J., Nolan, R., Venable, W., & Miller, M. (1998). Using deliberative inquiry to assess future needs for a graduate HRD curriculum: A case study. *11*(4), 76-90.

Spence, L. (2000). The same but different: Standardizing interview procedures in the multinational setting. *13*(2), 43-55.

Stamper, C., & VanDyne, L. (1999). Diversity at work: Do men and women differ in their organizational citizenship behavior? *12*(1), 59-76.

*Stolovitch, H., & Yapi, A. (1997). Use of case study method to increase near and far transfer of learning. *10*(2), 64-82.

*Sugrue, B., Rietz, T., & Hansen, S. (1999). Distance learning: Relationships among class size, instructor location, student perceptions and performance. *12*(3), 44-57.

*Tesoro, F. (1998). Implementing an ROI measurement process at Dell Computer. *11*(4), 103-114.

Thatcher, S. (1999). The contextual importance of diversity: The impact of relational demography and team diversity on individual performance and satisfaction. *12*(1), 97-112.

Tovar, M., Gagnon, F., & Schmid, R. (1997). Development of a consultation profile of interventions perceived as successful by human performance technology consultants. *10*(3), 67-83.

Twitchell, S., Holton, E., III, & Trott, J., Jr. (2000). Technical training evaluation practices in the United States. *13*(3), 84-109.

*Wagner, K., & Bailey, J. (1998). The effects of a monetary incentive system on work performance of mental health counselors. *11*(3), 64-78.

*Willis, V. (1998). Action learning: Design features and outcomes at Georgia State University. *11*(2), 34-47.

*Yelon, S., Reznich, C., & Sleight, D. (1997). Medical fellows tell stories of application: A grounded theory on the dynamics of transfer. *10*(2), 134-155.

*Yiu, L. (1998). Use of action learning as a vehicle for capacity building in China. *11*(1), 129-148.

*Yorks, L., O'Neil, J., Marsick, V., Lamm, S., Kolodny, R., & Nilson, G. (1998). Transfer of learning from an action reflection learning (TM) program. *11*(1), 59-73.