USING IMPORTANCE-PERFORMANCE ANALYSIS TO EVALUATE TRAINING

Jason M. Siniscalchi | Edward K. Beale | Ashley Fortuna

The importance-performance analysis (IPA) is a tool that can provide timely and usable feedback to improve training. IPA measures the gaps between the importance and how "good" (performance) a class is perceived by a student and is presented on a 2x2 matrix. The quadrant in which data land in this matrix aids in determining potential future action. This article focuses on integrating IPA into Kirkpatrick's (1959) four-level framework of evaluation. Examples and suggestions for improving training are discussed.

IMPORTANCE-PERFORMANCE ANALYSIS (IPA) is a common technique that is used to understand customer satisfaction and service quality. As an evaluation tool, IPA graphically depicts the gap between a current perceived state or condition and a future expected outcome. This visualization tool can then be used to quickly and effectively improve decision making and begin further investigation of root causes. Initially published by Martilla and James (1977) to improve marketing strategies, it has since been used in other management fields including recreation (Hammitt, Bixler, & Noe, 1996), tourism and leisure services (Oh, 2001), and food services (Sampson & Show-alter, 1999), among numerous other disciplines. At the United States Coast Guard Academy's Leadership Devel_____ use IPA to improve leadership training by using the IPA output to help better understand if our instruction, instructors, and learning outcomes are effectively "hitting the mark," and thereby improving our course curriculum and leadership training.

TRAINING EVALUATION

The LDC offers 23 courses, with an annual throughput of nearly 6,000 students, each of which are occasionally evaluated based on Kirkpatrick's (1959) evaluation framework. Kirkpatrick proposed four levels of evaluation:

- Level 1—Reaction: Captures the student's satisfaction.
- Level 2—Learning: Assesses the student's ability to demonstrate mastery of terminal performance objectives (TPO) in the training environment.
- Level 3—Behavior: Measures extent of actual on-the-job performance.
- Level 4—Results: Determines benefit to the organization.

Level 1 and 3 evaluations are the focus of this article. Level 1, Reaction, measures the student's perceptions of the course content, materials, learning environment, and instructor's performance. The purpose is to capture the student's perspective of the training as well as to enhance learning transfer by affording the students an opportunity for input on their training experience.

Level 3, Behavior, measures the graduate's "on the job" application of learned knowledge or skills in the actual working environment. Typically, a good rule of thumb is to schedule the evaluation approximately 3 to 6 months after the training is completed. Additionally, these evaluations are unique in that they require the willing participation of both the graduates and their first-line supervisors (Kirkpatrick & Kirkpatrick, 2006).

Despite the importance of evaluating training for effectiveness, literature has shown that evaluation of training programs is often incomplete or missing (Fullard,

2007). Adapting the IPA approach as a tool to present level 1 and 3 evaluations can help overcome this limitation.

IPA DEFINED

Nale, Raunch, Wathen, and Barr (2000) suggest that studies directed at improving higher education outcomes have a drawback: they focus exclusively either on importance or on performance. To alleviate this concern, the two factors can be combined (McPhillip, 2001; Nale et al.). Traditional IPA measures the gaps between the perceived importance of an attribute and how "good" (performance) the attribute is perceived to be by a user or customer. Scores of central distribution (means or medians) on each attribute's importance and performance are calculated, and its coordinates are graphed to form a two-dimensional matrix called an action grid (Blake, Shrader, & James, 1978). Each attribute will fall into one of four quadrants defined by Ortinau, Bush, Bush, and Twible (1989). Data landing in the upper right quadrant (high importance/high performance), defined as "keep up the good work," suggests that the current conditions and expected outcomes are being met—these are strengths. The upper and lower left quadrants (low importance/ high performance and low importance/low performance), labeled "possible overkill" and "low priority," suggest that the attribute may be superfluous due to low importance—these attributes may be candidates for discontinuation. Conversely, the lower right quadrant (high importance/low performance) is identified as "concentrate here." Data in this quadrant indicate importance outweighs ability, and these attributes present opportunities for corrective action.

APPLICATIONS OF IPA IN EDUCATION AND TRAINING

IPA has been used effectively in other educational institutions (Alberty & Mihalik, 1989; Attarian, 1995; Duke, 2002; Nale et al., 2000; Ortinau et al., 1989). Our adaptation of the IPA approach seeks to continue these efforts by using IPA as a method to close the loop between curriculum design, instructor/course feedback in class (level 1), and graduate performance of learned tasks on the job (level 3).

IPA FOR LEVEL 1 EVALUATIONS

Level 1 evaluations are used to measure a student's reaction to the course material and instructor performance. Typical analyses compare both course content and instructor performance as separate dimensions. It is our contention, however, that the two are related measures. By graphing this relationship, we can gain new insights into

Despite the importance of evaluating training for effectiveness, literature has shown that evaluation of training programs is often incomplete or missing (Fullard, 2007). Adapting the IPA approach as a tool to present level 1 and 3 evaluations can help overcome this limitation.

the interaction between the instructor's delivery of material and the usability of the information presented. To make this comparison, we have adjusted the IP axes to relate course content ("importance") to instructor performance ("ability"). Specifically, we ask students to evaluate the following items on a 5-point Likert-type scale: To what extent do they agree the information they learned will help them on the job and to what extent do they agree that the instructor demonstrated knowledge of the material presented. The quadrants have been renamed (from upper right, clockwise) as follows: "keep up the good work," "instructor training needed," "revisit curriculum and/or lesson plan," and "focus on objectives" (see Figure 1).

A check of validity of this matrix can be seen in student comments regarding both the instructor and the course content for each block of instruction. Interpreting the data in Figure 2, which represents a particular module of instruction (number) and instructor who taught the module (letter), we see comments directly relate to each quadrant on the IP matrix.

Focus on Objectives (8Q)

- "I'm beginning to see a trend of side bar conversations going on longer with the instructor. While many of these conversations would likely be beneficial to much of the class, they are being conducted not involving the entire class and I think the class gets off track."

- "[The instructor] needed to stay more on subject. He was a nice and energetic person but jumped off subject and was hard to follow most of the time."

Revisit Curriculum or Lesson Plan (14G, 3G)

- "[The instructor] is very nice and versed, but, perhaps if the material was a bit more updated that would help. Using real world examples [would help] as well."

- "He had to skim over a lot of material very quickly and [told us he had to get] ready for this on short notice . . . which is more the fault of the course trying to squeeze everything in."

- "Could have had an explanation beforehand of what the class would cover instead of jumping and giving info in sections and leaving us to figure it out in the end."

Instructor Training Needed (1P)

- "I believe [the instructor], albeit a wonderful lady, didn't know all the info. She had to keep relying on [another instructor] for the answers. The two combined made a great instructor."

- "[The instructor] did a fine job instructing, but his style differs somewhat from the other instructors to date. It just took a little while for me to adjust to his style . . . it was hard to follow."

IPA WITH LEVEL 3 EVALUATIONS

Level 3 evaluations assess the degree to which classroom knowledge transferred to the actual job performance. Incorporating level 3 evaluations with IPA is a straightforward way to help compare graduate ability to an important terminal performance objective (TPO), defined in training. For each TPO, we ask both the graduates and their supervisors to evaluate the following items on a 5-point scale: Rate the graduate's ability to perform the job task from 1 = *poor ability* to 5 = *expert ability,* and rate how important this TPO is to the overall job (or, in the case of the military, mission) from 1 = *not important* to 5 = *very important*. The original quadrant labels are appropriate (from upper right, clockwise) "keep up the good work," "concentrate here," "low priority," and "possible overkill" (see Figure 3). The supervisors' assessment of importance and the graduates' perception of ability (based on what they learned in class) is plotted to form the Importance-Performance matrix.

An example of the level 3 IPA is found in Figure 4, which displays the supervisor's assessment of graduate ability and TPO importance for one class at the LDC. Numbers on the matrix refer to individual TPOs. In this figure, we see that most TPOs fall in the upper right quadrant ("keep up the good work"), suggesting a majority of the TPOs taught are on the mark. However, two TPOs fall into the upper left quadrant ("possible overkill"), suggesting graduates perform well on TPOs identified as less important relative to the other TPOs taught. Additionally, one TPO is not on the matrix because no graduate indicated that he or she performs the TPO. Though these data are from one session of one course, if this trend continues in future classes and further investigation into these modules is warranted, it may result in a modification of the curricula.
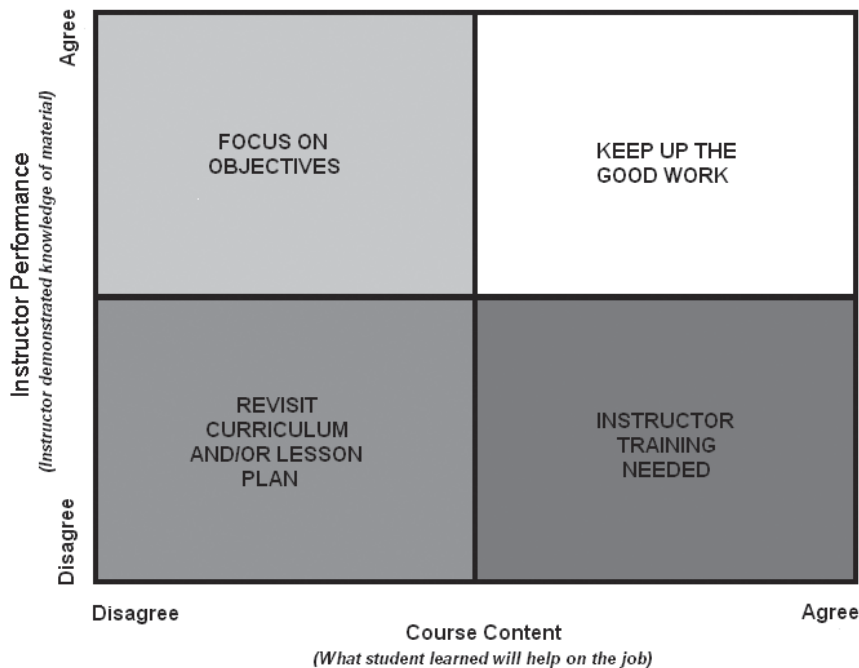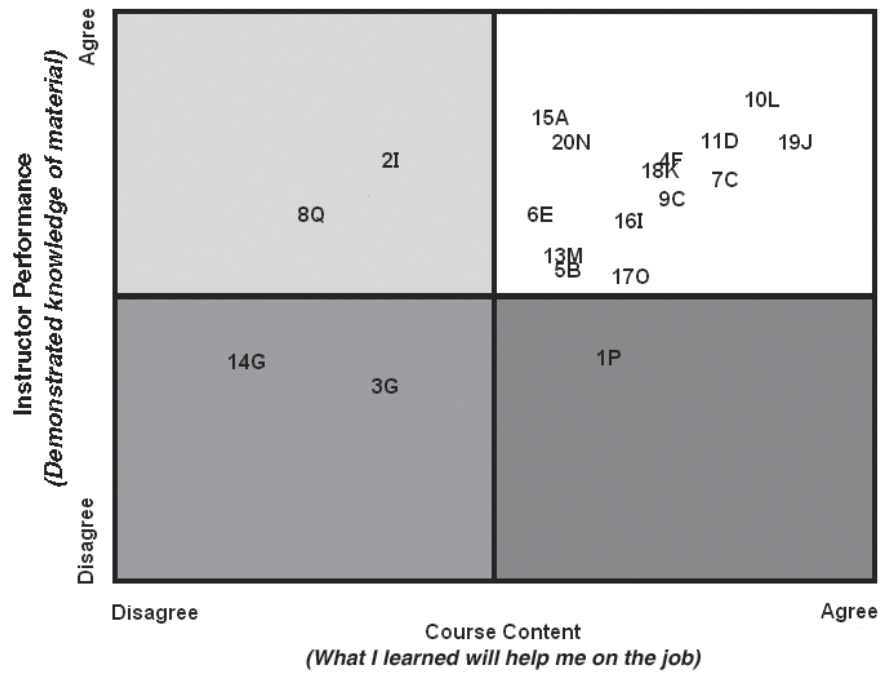

FIGURE 1. LEVEL 1 EVALUATION IP MATRIX

Agree

Instructor Performance
(Demonstrated knowledge of material)

15A
10L
20N   11D   19J
2I   18K 4F
9C   7C
8Q   6E   16I
13M
5B   17O

Disagree

14G   1P
3G

Disagree   Course Content   Agree
(What I learned will help me on the job)

Note. Data labels represent the block of instruction or lesson (number) and the instructor (letter).

FIGURE 2. EXAMPLE OF LEVEL 1 IPA MATRIX

IPA AND CURRICULUM DEVELOPMENT

The analysis of evaluation data is most valuable to an instructional program when the results serve to close the loop in the larger curriculum and lesson design process.

Adhering to the ADDIE instructional design model, a mid-stream curriculum review should begin with an analysis phase, and historical evaluation data should be an integral part of the analysis. Here is where the graphical representation in an IPA can add the most value.
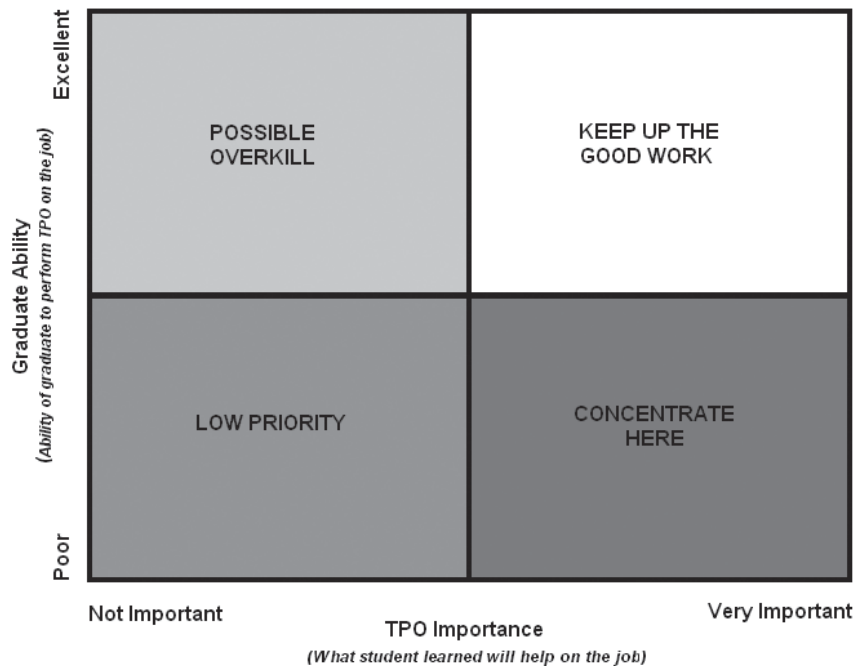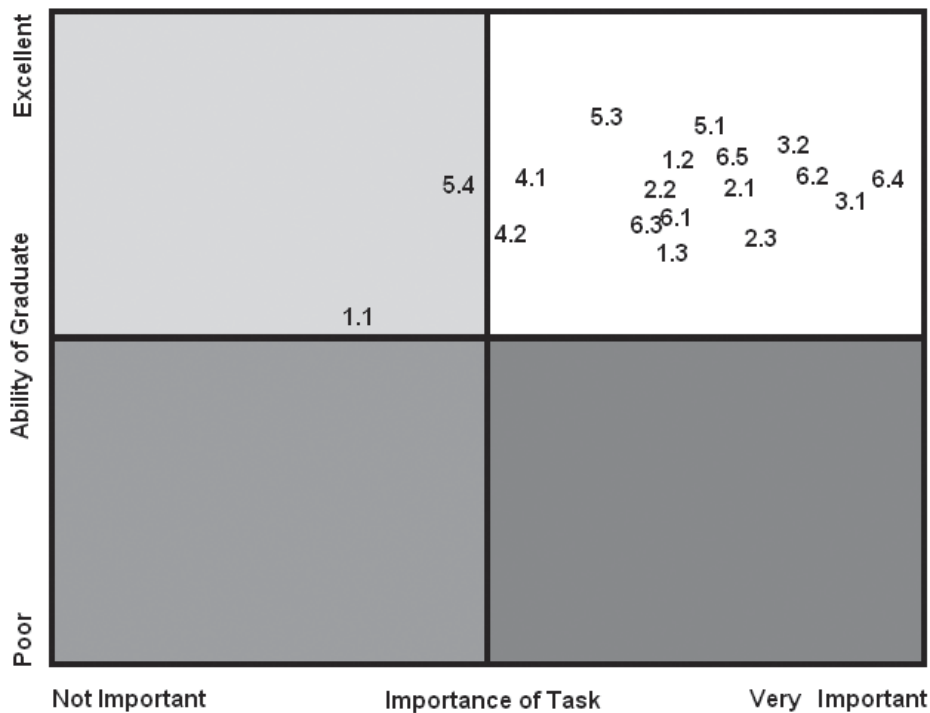

Excellent

Graduate Ability
(Ability of graduate to perform TPO on the job)

POSSIBLE OVERKILL   KEEP UP THE GOOD WORK

LOW PRIORITY   CONCENTRATE HERE

Poor

Not Important   TPO Importance   Very Important
(What student learned will help on the job)

FIGURE 3. EVALUATION IP MATRIX

Note. Numbers indicate TPOs. No graduate performed TPO 5.2.

FIGURE 4. EXAMPLE OF LEVEL 3 IPA

Those elements of the instructional program that plot low either on the importance axis or on the performance axis (or both) are likely to need the most review.

For example, a level 3 (transfer) evaluation of a technical lesson may reveal that the students have a high level of skill on the job but that the skill is not at all important. Should you really be wasting instructional time teaching this lesson? Perhaps. Training the skill may be mandated by a regulation, as is often the case with compliance lessons such as safety, ethics, or workplace climate. However, in the absence of a mandatory requirement, your data have lent credence to the decision to remove the lesson from your program.

Conversely, you may discover that a task plots as extremely important to the job, but the students have a low level of skill completing the task. Your data have revealed a starting point for additional inquiry. Is student skill low because your training is inadequate or because students do not have enough opportunities to practice? Each reason will lead to a different instructional approach in your revised curriculum.

Keep in mind that the data may indicate a healthy program that should *not* be changed, as evidenced by most or all lessons plotting high in both importance and performance. Such data can help counter pressure to make an

arbitrary change because your program is clearly training the most important skills and training them well. Why change the recipe? Plotting your level 1 and level 3 data as a function of importance and performance allows you to easily identify those lessons that are working, and focus on the lessons that may need revision or removal.

CONCLUDING THOUGHTS

IPA is an established and effective evaluation tool that is easy to apply. It provides a visualization of data that affords immediate feedback and can be used to facilitate change in areas of concern. In a training environment, IPA can help by not only improving instructor performance by indicating areas of instruction not "hitting the mark" but also modifying a curriculum so future students can learn and master objectives directly applicable to their job performance. It should be noted that using a survey to derive level 3 information provides, at best, an indicator of a potential problem. A true level 3 evaluation would require an in-person audit of task ability, not just a perception of ability. IPA, however, can identify potential areas of concern in a curriculum, particularly when the assessment incorporates data from multiple sessions of a class. Finally, an IPA can help close the loop between

existing instruction and a systematic review of instructional curriculum. ▲

References

Alberty, S., & Mihalik, B.J. (1989). The use of importance-performance analysis as an evaluative technique on adult education. *Evaluation Review, 13*(1), 33–44.

Attarian, A. (1995). Using importance-performance analysis to evaluate teaching effectiveness. In R. Koesler & R. Watters (Eds.), *Proceedings of the 1995 International Conference on Outdoor Recreation and Education* (pp. 145–150).

Blake, B.F., Schrader, L.F., & James, W. (1978). New tools for marketing research: The action grid. *Feedstuffs, 50*(19), 38–39.

Duke, C.R. (2002). Learning outcomes: Comparing student perceptions of skill level and importance. *Journal of Marketing Education, 34*(3), 203–217.

Fullard, F. (2007). A model to evaluate the effectiveness of enterprise training programmes. *The International Entrepreneurship and Management Journal, 3*(3), 263–276.

Hammitt, W.E., Bixler, R.D., & Noe, F.P. (1996). Going beyond importance-performance analysis to analyze the observance-influence of park impacts. *Journal of Park and Recreation Administration, 14*(1), 45–62.

Kirkpatrick, D., & Kirkpatrick, P. (2006). *Evaluating training programs* (3rd ed.). San Francisco: Berrett-Koehler.

Kirkpatrick, D.L. (1959). Techniques for evaluating training programs. *Journal of American Society for Training and Development, 13*, 11–12.

Martilla, J.A., & James, J.C. (1977). Importance-performance analysis. *Journal of Marketing, 10*(1), 13–22.

McPhillip, J. (2001). Case studies in job analysis and training evaluation. *International Journal of Training and Development, 5*(4), 283–289.

Nale, A.R., Rauch, D.A., Wathen, S.A., & Barr, P.B. (2000). An exploratory look at the use of importance-performance analysis as a curricular assessment tool in a school of business. *Journal of Workplace Learning: Employee Counseling Today, 12*(4), 139–145.

Oh, H. (2001). Revisiting importance-performance analysis. *Tourism Management, 22*(6), 617–27.

Ortinau, D.J., Bush, R.P., Bush, A., & Twible, J.L. (1989). The use of importance-performance analysis for improving the quality of marketing education: Interpreting faculty-course evaluations. *Journal of Marketing Education, 11*(2), 78–86.

Sampson, S.E., & Showalter, M.J. (1999). The performance-importance response function: Observations and implications. *The Service Industries Journal, 19*(3), 1–25.

JASON M. SINISCALCHI, PhD, is a research social scientist at the U.S. Coast Guard Academy's Leadership Development Center. He is responsible for implementing and analyzing all resident and nonresident training evaluations of the Leadership Development Center. He earned his doctorate from West Virginia University where he conducted research on improving program evaluation reports through mixed method research designs and integrating data visualization techniques to create "usable knowledge." His other research interests include integrating GIS into social science research and investigating the role of the setting in performance improvement. He may be reached at Jason.M.Siniscalchi@uscg.mil.

EDWARD K. BEALE, CDR, USCG, is the training branch chief at the U.S. Coast Guard Academy's Leadership Development Center. He earned an undergraduate degree in civil engineering from the U.S. Coast Guard Academy and a graduate degree from San Diego State University in educational technology. He has authored several papers addressing human performance and process improvement. He is a rescue helicopter pilot with operational experience in the polar regions and in Hurricane Katrina. He may be reached at Edward.K.Beale@uscg.mil.

ASHLEY FORTUNA is an intern at the U.S. Coast Guard Academy's Leadership Development Center. Currently a senior at Sacred Heart University in Fairfield, Connecticut, she is majoring in business administration and concentrating in international business. She has interests in human resources management and legal studies and is considering pursuing a law degree in International Legal Studies. She may be reached at Fortunaa@sacredheart.edu.