

# The Feds Lead the Way in Making Training Evaluations More Effective

By James D. Kirkpatrick and Wendy K. Kirkpatrick

U.S. government agencies are taking heed to the recommendations outlined in the 2009 ASTD report, *The Value of Evaluation*.



LISTEN TO THIS FEATURE  
at [www.astd.org/TDpodcasts](http://www.astd.org/TDpodcasts)



.....

In 2009, ASTD published *The Value of Evaluation: Making Training Evaluations More Effective*, a report that revealed the results of a study that queried organizations regarding how well training evaluation was meeting their business needs. Responses to the 26 questions led to disturbing conclusions. “The pursuit of excellent learning evaluations continues, but so far few organizations think they’ve mastered them,” the report states. “Only about one-quarter of respondents ... agreed that their organization got a solid ‘bang for the buck’ from its training evaluation efforts.”

### **Priorities for improvement**

The five most urgent recommendations resulting from the study (reordered into the sequence in which they generally occur) are to:

- ensure that learning programs positively influence employee behaviors
- ultimately improve overall business results
- calculate the effect of learning on important business results
- gauge the return-on-investment of learning programs
- demonstrate to others in the organization the value of the learning function.

Anecdotally, most training professionals will admit that, even three years after the study, they continue to struggle with these five priorities. They remain mired in the measurement of “effective training,” which relates to Kirkpatrick Levels 1 and 2, and cannot seem to get to “training effectiveness,” or Kirkpatrick Levels 3 and 4, which relate to the five priorities for improvement (see Figure 1).

Therefore, it is notable that organizations in the public sector are making significant strides in each priority.

### **Progress is being made**

U.S. government agencies are moving beyond a training event focus and are realizing the benefits of measuring training effectiveness—defined as the degree to which training and subsequent reinforcement help workers to perform specific tasks that contribute to organizational objectives. Real evidence exists that traditional training events are being replaced with holistic learning and performance support processes.

Here are the stories of three government agencies, which should inspire and assist all training professionals to create and demonstrate organizational value in their work.

## U.S. Office of Personnel Management (OPM)

OPM is serious about training contributing to the missions of the more than 200 federal agencies that serve U.S. citizens at home and abroad. OPM's mission is to recruit, retain, and honor a world-class workforce to serve the American people. It is the human resources function of the federal government, if you will.

Part of OPM's function is to create and implement federal regulations for training and development. OPM has paralleled ASTD's recommendations to reinvent training evaluation by drafting regulations used as cornerstones for its *Training Evaluation Field Guide* published in 2011.

Two of the key regulations corroborate the five ASTD recommendations. The first (5 CFR Sec. 410) states that "agencies must develop and implement a process to evaluate training and development program impact in terms of: a) learning, b)

employee performance, c) work environment, [and] d) contribution to mission accomplishment, and the results of the evaluation must reflect a positive contribution to mission accomplishment."

The second regulation (5 CFR Sec. 250) states that "agencies must evaluate each program or plan established, operated, or maintained under subsection (a) with respect to accomplishing specific performance plans and strategic goals and modify such programs or plan as needed to accomplish such plans and goals."

OPM regulations provide the guidelines for all U.S. federal government agencies to create and demonstrate organizational value with their training investments. The field guide provides a road map. (Read more about the resources OPM provides in the sidebar on page 59.)

## The Federal Law Enforcement Training Accreditation Board/ Office of Accreditation (FLETA/OA)

Billy McLeod is a program manager for FLETA, under the umbrella of the U.S. Department of Homeland Security. McLeod's work is law enforcement sensitive, so he could share only a high-level

overview of the agency's shift to a full four-level evaluation methodology and related accomplishments.

The federal law enforcement community ensures that it provides the most current, relevant training available through continual evaluation at all levels. Many of the law enforcement agencies within the U.S. Department of Homeland Security, Department of Justice, Department of Defense, and numerous other departments have demonstrated this professionalism by seeking accreditation through FLETA. Evaluation is specifically addressed within four of the 52 standards that the training academies must demonstrate to earn program accreditation.

Three of the standards that address training evaluation are:

- 3.54: The applicant conducts, compiles, and reviews student reaction surveys (Level 1 of the Kirkpatrick model or an equivalent) to identify opportunities to improve the program, instruction, support, and administrative elements of the training received.
- 3.55: The applicant reviews the results of the program's student examinations (Level 2 of the Kirkpatrick model or an equivalent) to identify gaps in instruction, student materials, or deficiencies in test-item construction.
- 3.56: The applicant gathers and reviews feedback (Level 3 of the Kirkpatrick model or an equivalent) from graduates and their supervisors to evaluate the effectiveness of the program.

Adherence to these and many other standards has allowed the participating agencies to make significant improvements that have delivered numerous benefits throughout their organizations.

FLETA also practices what it preaches. McLeod has successfully completed a Level 4 evaluation that examined both quantitative and qualitative data as part of a case study of a Federal Law Enforcement Training Center training program.

## Figure 1 | The Kirkpatrick Model

<b>Effective Training</b>	<b>LEVEL 1</b>	To what degree participants react favorably to the learning event
	<b>LEVEL 2</b>	To what degree participants acquire the intended knowledge, skills, and attitudes based on their participation in the learning event
<b>Training Effectiveness</b>	<b>LEVEL 3</b>	To what degree participants apply what they learned during training when they are back on the job
	<b>LEVEL 4</b>	To what degree targeted outcomes occur as a result of learning event(s) and subsequent reinforcement.



## OPM Does Much More Than Draft Policies and Regulations

In post-program surveys, students ranked their preparedness to perform individual skills and lessons taught within the program at 92 percent to 96 percent. Equally impressive is that supervisors ranked student preparedness at 92 percent to 93 percent. Students and their supervisors were asked, “Can you identify any successes that this training prepared you for?” and “Have you applied skills learned in this training program to other areas of your operations?” Responses showed that learning transfer had occurred and that agencies had realized numerous gains.

This four-level evaluation demonstrates that the program is meeting the expectations of both the organization and its customers, and also is meeting multiple national security strategies. The program brought about change both with the student and within the agency, and has even positively transformed other areas of operations.

Additional Level 4 evaluations that measure multiple levels of operations within a law enforcement component are ongoing within the FLETA community.

### Internal Revenue Service (IRS)

In 2008, Rob LaVanway, then working with the Large and Mid-Size Business Division of the IRS, received marching orders to “hire a substantial number of new mid-career employees while minimizing field resources for training.”

LaVanway and his team first identified Level 4 organizational goals, and then worked backward through Levels 3, 2, and 1 to develop a flexible blended learning and performance strategy by:

- aligning objectives to level 4 business results
- leveraging technology-enabled learning
- creating consistency
- reducing use of instructors and coaches
- maximizing existing learner knowledge, skills, and abilities.

The legacy training program was replaced by one built from the Large and

Mid-Size Business Division’s operational goals. This flexible, blended learning model was augmented by targeted activities before and after training designed to maximize on-the-job application and results.

This new approach to training and development required substantial change management and socialization efforts. It was implemented as an inclusive process that was open to and inviting of new ideas and best practices from the frontline implementers.

The solution was presented as a model that would be most effective when fitted for the specific needs of the trainee. Therefore, the new employees’ managers and coaches were encouraged to consult with the training department as much as necessary, but ultimately were empowered to make decisions that balanced employee training needs, customer satisfaction, and specific business plan goals. In return, training staff asked for and received support in implementing the evaluation methodology.

One IRS business unit had experienced a 40 percent attrition rate during a four-year period for one of its most technically intensive positions. One might anticipate that a successful training program for a position with such a steep learning curve for new recruits might score well at Kirkpatrick Levels 1 and 2 but would suffer at Levels 3 and 4, consistent with the high attrition rate. Instead, by aligning the program to Level 4 results, concentrating on pretraining preparation and required drivers, and using a business partnership model, Level 3 data revealed that new recruits performed better than journeyworkers during the same measurement period (see Figure 2).

Level 3 data also revealed that journeyworker performance improved during the same measurement period. Though this is likely a factor of multiple inputs, it may suggest that learning transfer occurred. New-hire scores in Year 1 appear to indicate behavior transfer, as evidenced by increasing journeyworker scores by Year 3. When Year 1 new-hires

Sure, the U.S. Office of Personnel Management (OPM) drafted the regulations that laid the groundwork for a new world training evaluation methodology. It also led the charge, along with Kirkpatrick Partners and Booz Allen Hamilton, to develop the *Federal Training Evaluation Field Guide*. Now the agency is just as passionate about seeing that it is implemented across the more than 200 federal agencies.

Here are some implementation tips and resources gleaned from OPM documents and conversations.

In training (or at least federal training), establishing a direct cause-and-effect relationship between training and the business result is quite difficult. Agencies must consider other external factors (for example, learner motivation, supervisory support, and opportunity to apply the learning) that influence the accomplishment of the program result.

*The Guide to Strategically Planning Training and Measuring Results* is available to assist agencies in taking a strategic approach to training. It is being updated to reflect the new requirements of the 2010 Government Performance Results and Accountability Modernization Act as they relate to training, but the current guide reflecting the requirements of the 1993 Government Performance Results and Accountability Act remains quite useful.

A governmentwide training and development wiki with evaluation resources, free training, and other relevant information for agencies is live ([www.opm.gov/wiki/training/Training-Evaluation/Print.aspx](http://www.opm.gov/wiki/training/Training-Evaluation/Print.aspx)). This augments the *Federal Training Evaluation Field Guide* and provides a collaboration venue.

OPM continues to offer technical assistance through face-to-face agency presentations, webcasts, forums, roundtables, web chats, and other information-sharing and instructional techniques.

Figure 2 | Level 3 Data of an IRS Business Unit



became journeyworkers by Year 3, overall workforce scores began to trend up.

In 2009, the IRS presented how the Large and Mid-Size Business Division's initiative employed the Kirkpatrick best practice of "starting with the end in mind." This new-recruit training strategy measures training effectiveness at all four Kirkpatrick levels, aligning them by terminal and enabling objectives comprised of the critical success factors that ultimately deliver the agency's desired business results.

The four levels form a "chain of evidence" for the IRS that enables it to demonstrate (by a preponderance of evidence) the effectiveness of the new-recruit training in supporting the organization's successful business results. By aligning objectives with critical Level 3 behaviors and Level 4 results, IRS training strategy and curriculum improvement is now a programmatic and iterative process tied to specific business strategies (see Figure 3).

The IRS does not compute ROI, but it does align training to a balanced measurement system and forms a chain of evidence to demonstrate training's effectiveness in contributing to mission accomplishment. This methodology is in line with the Kirkpatrick Model, ROE (return on expectations), and the recent OPM regulations.

The three balanced criteria for ultimate Level 4 success were achieved while cutting training costs by 50 percent. This counters the common and detrimental belief that Level 3 is expensive to measure, and Level 4 nearly impossible.

In addition to regular program updates, the following methods also were used to demonstrate program value.

- Executives were briefed regarding overall recruit training effectiveness (Levels 1 through 4 chain

All U.S. residents benefit from more efficient government administration when U.S. government agencies create training effectiveness.



Figure 3 | Level 4 Business Results for the IRS

Balanced Measures	2008	2009	2010	2011
Employee engagement	78%	82%	83%	84%
Customer satisfaction	82%	79%	81%	81%
Business results	Green	Green	Green	Green

Follow the lead of these agencies, and help your organization and clients to view the value of training as the accomplishment of Kirkpatrick Levels 3 and 4.

of evidence) for 2008 through 2012 (ongoing).

- The model is communicated as a prototype when new training strategies are implemented (ongoing).
- The model and results are shared with other IRS business units that request support in implementing the same methodology for their organizations.

The following collateral benefits of the new approach also were presented to stakeholders.

- Substantial portions of HR-intensive training were converted from in-person to technologically enabled.
- Updated training review processes that also serve as Levels 2 and 3 evaluations to unburden frontline managers were used.
- Frontline managers and coaches were relieved from delivering numerous workshops through use of centralized virtual conferencing and performance support job aids as alternatives to instructor-led learning.

### Following the U.S. government's lead

All U.S. residents benefit from more efficient government administration when U.S. government agencies create training effectiveness. The trailblazers in this article are not finished with their journey, and admittedly have not gotten

every detail correct along the way. The important factor is that they have started a necessary movement. They fought the temptation to stay within the comforts of measuring only Kirkpatrick Levels 1 and 2, and have taken action to meet OPM regulations and the 2009 *Value of Evaluation* report recommendations.

Follow the lead of these agencies, and help your organization and clients to view the value of training as the accomplishment of Kirkpatrick Levels 3 and 4. OPM has shown that even in a position of influence without absolute control, good progress can be made. McLeod has shown that even in high-security, confidential situations, data can be gathered to show a chain of evidence. And LaVanway has shown that measuring Levels 3 and 4 is neither difficult nor expensive when they are used to set objectives at the beginning of an initiative.

The fear of being unable to prove the connection between training efforts and Level 4 results often is cited as a reason to not even try to measure training effectiveness. The conclusion of the *Value of Evaluation* report provides sage advice: "One possible takeaway is that companies should not try to prove that a learning experience affects results but rather show that, given the preponderance of evidence, it very likely does."

The writing is on the wall for the entire training industry: create and demonstrate business and mission value at Level 4 or

risk becoming a nonfactor. If an entity as complex as the U.S. federal government can figure out how to do it, so can you.

The following individuals contributed to this article: Rob LaVanway, technical training team manager, International, Internal Revenue Service; and Billy McLeod, program manager, Federal Law Enforcement Training Accreditation Board/Office of Accreditation.

**James D. Kirkpatrick** is a senior consultant for Kirkpatrick Partners; [information@kirkpatrickpartners.com](mailto:information@kirkpatrickpartners.com). **Wendy K. Kirkpatrick** is president of Kirkpatrick Partners; [information@kirkpatrickpartners.com](mailto:information@kirkpatrickpartners.com).

#### INTERESTED IN ORDERING E-PRINTS?

Would a digital version of this article be a great fit for your next course, presentation, or event? Are you interested in e-prints of several *T+D* articles on a specific topic? Visit [www.astd.org/TD](http://www.astd.org/TD), and click on "About T+D," for more information.