Invited Reaction: Reaction to Holton Article

Donald L. Kirkpatrick

I didn't know whether to laugh or cry when I read the title to Holton's article, "The Flawed Four-Level Evaluation Model." After I read the article, I still didn't know how to feel. But when I thought about how Holton had "proved"—through various research papers—that the Kirkpatrick model wasn't really a model at all but only a "taxonomy," I decided not to do either. I will admit that I was a little upset when he listed the six components of a model described by Klimoski (1991) and Dubin (1996) and then stated that "the four-level model does not meet any of these criteria." He might at least have said that it met the second criterion of "relationships between the units" because my model (or my taxonomy, if you prefer) does show the relationships among the four levels.

I admit that Holton's article is a scholarly work; Certainly, cites to other articles are plentiful. Many of them have nothing to do with evaluation, but the list is impressive.

The funny thing is that I personally have never called my framework "a model." Someone else described it that way. For example, in a case study presented in Evaluating Training Programs: The Four Levels (Kirkpatrick, 1994), Dave Basarab, head of the evaluation department at Motorola stated, "Motorola University has adopted the Kirkpatrick model for training evaluation" throughout the world. In another case study presented in the book, the authors from CIGNA Corporation called my model "the CIGNA CMD&T Impact Model." The case study by Patrick O'Hara of First Union National Bank referred to it as "the Four-Level Kirkpatrick Evaluation Model." And Eric Freitag of Intel Corporation wrote an article for the book entitled "Implementing the Kirkpatrick Model as an Up-Front Analysis and Evaluation Tool."

My purpose in offering these illustrations is to demonstrate that, in the real world where training evaluation takes place, the word "model" is commonly used to describe a systematic way of doing something. It may or may not meet the six criteria listed by Klimoski. Personally, I don't care whether my work is called a model or a taxonomy as long as it helps to clarify the meaning of...
evaluation in simple terms and offers guidelines and suggestions on how to accomplish an evaluation. On second thought, I am glad it is not referred to as a taxonomy because if it were trainers wouldn't know what it meant. The word “model” seems to communicate that it is something to be used as a helpful guide.

I realize that the Human Resource Development Quarterly is a scholarly journal that does not publish simple, practical articles. As a former full professor with a Ph.D., I also realize that publishing scholarly articles is required for promotion and recognition in academic circles. I realize too that those who use my model will probably continue to read the Journal of the American Society of Training and Development, where the emphasis is on helping the audience rather than on demonstrating the amount of scholarly research done.

Yet I admit I was a little upset by the word “flawed” used in Holton's title. My four-level model has been quoted and used all over the world. Training professionals find it helpful in doing something about evaluation. I have conducted sessions on it at professional conferences in the United States, Venezuela, Mexico, and Singapore. Most recently, in November 1995, I conducted four days of programming for the Arabian American Oil Company in Saudi Arabia.

Indeed, the Holton article tried to tear down the four-level model without giving any credit to the concepts it depicts. For example, Holton used the term reaction, but not as a level. And the “model” he described used the term learning, which is the second level of the Kirkpatrick model. It is interesting that it did not use the term behavior but instead a very similar term, performance. And, finally, it avoided the word results as the fourth level by substituting the words organizational results. It seems to me that if my model was so flawed, Holton would not have included so much of it in his concept of a true model.

Finally, in his conclusions, Holton stated that “the model presented here [in his article] is an initial step in the development and testing of a true model of HRD evaluation.” He ended by stating that, “If HRD is to grow as a discipline and as a profession, it is imperative that researchers work deliberately to develop a more integrative and testable model.” I note that he works at a vocational school. I don’t know what qualifications or experience he has with the HRD profession. In contrast, I am a past president of the American Society for Training and Development (ASTD) as well as an author, speaker, and consultant in the field. I regularly conduct workshops at the National Convention of ASTD on the subject of evaluation. These sessions always seem to draw a full house of HRD professionals who are looking for help in evaluating training programs. I only hope that my model—or taxonomy—continues to be of help to the HRD professionals who are more interested in practical ideas than in scholarly research.
References


Donald L. Kirkpatrick is professor emeritus of management, University of Wisconsin.