

Kirkpatrick's Four Levels of Training Evaluation in Detail

This grid illustrates the Kirkpatrick's structure detail, and particularly the modern-day interpretation of the Kirkpatrick learning evaluation model, usage, implications, and examples of tools and methods.

EVALUATION TYPE	EVALUATION DESCRIPTION AND CHARACTERISTICS	EXAMPLES OF EVALUATION TOOLS AND METHODS	RELEVANCE AND PRACTICABILITY
<p>LEVEL 1 REACTION</p>	<ul style="list-style-type: none"> ▪ reaction evaluation is how the delegates felt, and their personal reactions to the training or learning experience, for example: ▪ did the trainees like and enjoy the training? ▪ did they consider the training relevant? ▪ was it a good use of their time? ▪ did they like the venue, the style, timing, domestics, etc? ▪ level of participation ▪ ease and comfort of experience ▪ level of effort required to make the most of the learning ▪ perceived practicability and potential for applying the learning 	<ul style="list-style-type: none"> ▪ typically 'happy sheets' ▪ feedback forms based on subjective personal reaction to the training experience ▪ verbal reaction which can be noted and analyzed ▪ post-training surveys or questionnaires ▪ online evaluation or grading by delegates ▪ subsequent verbal or written reports given by delegates to managers back at their jobs 	<ul style="list-style-type: none"> ▪ can be done immediately the training ends ▪ very easy to obtain reaction feedback ▪ feedback is not expensive to gather or to analyze for groups ▪ important to know that people were not upset or disappointed ▪ important that people give a positive impression when relating their experience to others who might be deciding whether to experience same
<p>LEVEL 2 LEARNING</p>	<ul style="list-style-type: none"> ▪ learning evaluation is the measurement of the increase in knowledge or intellectual capability from before to after the learning experience: ▪ did the trainees learn what intended to be taught? ▪ did the trainee experience what was intended for them to experience? ▪ what is the extent of advancement or change in the trainees after the training, in the direction or area that was intended? 	<ul style="list-style-type: none"> ▪ typically assessments or tests before and after the training ▪ interview or observation can be used before and after although this is time-consuming and can be inconsistent ▪ methods of assessment need to be closely related to the aims of the learning ▪ measurement and analysis is possible and easy on a group scale ▪ reliable, clear scoring and measurements need to be established, so as to limit the risk of inconsistent assessment ▪ hard-copy, electronic, online or interview style assessments are all possible ▪ 	<ul style="list-style-type: none"> ▪ relatively simple to set up, but more investment and thought required than reaction evaluation ▪ highly relevant and clear-cut for certain training such as quantifiable or technical skills ▪ less easy for more complex learning such as attitudinal development, which is famously difficult to assess ▪ cost escalates if systems are poorly designed, which increases work required to measure and analyze

EVALUATION TYPE	EVALUATION DESCRIPTION AND CHARACTERISTICS	EXAMPLES OF EVALUATION TOOLS AND METHODS	RELEVANCE AND PRACTICABILITY
<p>LEVEL 3 BEHAVIOR</p>	<ul style="list-style-type: none"> ▪ behavior evaluation is the extent to which the trainees applied the learning and changed their behavior, and this can be immediately and several months after the training, depending on the situation: ▪ did the trainees put their learning into effect when back on the job? ▪ were the relevant skills and knowledge used ▪ was there noticeable and measurable change in the activity and performance of the trainees when back in their roles? ▪ was the change in behavior and new level of knowledge sustained? ▪ would the trainee be able to transfer their learning to another person? <p>is the trainee aware of their change in behavior, knowledge, skill level?</p>	<ul style="list-style-type: none"> ▪ observation and interview over time are required to assess change, relevance of change, and sustainability of change ▪ arbitrary snapshot assessments are not reliable because people change in different ways at different times ▪ assessments need to be subtle and ongoing, and then transferred to a suitable analysis tool ▪ assessments need to be designed to reduce subjective judgment of the observer or interviewer, which is a variable factor that can affect reliability and consistency of measurements ▪ the opinion of the trainee, which is a relevant indicator, is also subjective and unreliable, and so needs to be measured in a consistent defined way ▪ 360-degree feedback is useful method and need not be used before training, because respondents can make a judgment as to change after training, and this can be analyzed for groups of respondents and trainees ▪ assessments can be designed around relevant performance scenarios, and specific key performance indicators or criteria ▪ online and electronic assessments are more difficult to incorporate - assessments tend to be more successful when integrated within existing management and coaching protocols ▪ self-assessment can be useful, using carefully designed criteria and measurements 	<ul style="list-style-type: none"> ▪ measurement of behavior change is less easy to quantify and interpret than reaction and learning evaluation ▪ simple quick response systems unlikely to be adequate ▪ cooperation and skill of observers, typically line-managers, are important factors, and difficult to control ▪ management and analysis of ongoing subtle assessments are difficult, and virtually impossible without a well-designed system from the beginning ▪ evaluation of implementation and application is an extremely important assessment - there is little point in a good reaction and good increase in capability if nothing changes back in the job, therefore evaluation in this area is vital, albeit challenging ▪ behavior change evaluation is possible given good support and involvement from line managers or trainees, so it is helpful to involve them from the start, and to identify benefits for them, which links to the level 4 evaluation below

EVALUATION TYPE	EVALUATION DESCRIPTION AND CHARACTERISTICS	EXAMPLES OF EVALUATION TOOLS AND METHODS	RELEVANCE AND PRACTICABILITY
<p>LEVEL 4 RESULTS</p>	<ul style="list-style-type: none"> ▪ results evaluation is the effect on the business or environment resulting from the improved performance of the trainee - it is the acid test ▪ measures would typically be business or organizational key performance indicators, such as: ▪ volumes, values, percentages, timescales, return on investment, and other quantifiable aspects of organizational performance, for instance; numbers of complaints, staff turnover, attrition, failures, wastage, non-compliance, quality ratings, achievement of standards and accreditations, growth, retention, etc. 	<ul style="list-style-type: none"> ▪ it is possible that many of these measures are already in place via normal management systems and reporting ▪ the challenge is to identify which and how relate to the trainee's input and influence ▪ therefore it is important to identify and agree accountability and relevance with the trainee at the start of the training, so they understand what is to be measured ▪ this process overlays normal good management practice - it simply needs linking to the training input ▪ failure to link to training input type and timing will greatly reduce the ease by which results can be attributed to the training ▪ for senior people particularly, annual appraisals and ongoing agreement of key business objectives are integral to measuring business results derived from training 	<ul style="list-style-type: none"> ▪ individually, results evaluation is not particularly difficult; across an entire organization it becomes very much more challenging, not least because of the reliance on line-management, and the frequency and scale of changing structures, responsibilities and roles, which complicates the process of attributing clear accountability ▪ also, external factors greatly affect organizational and business performance, which cloud the true cause of good or poor results

Since Kirkpatrick established his original model, other theorists (for example Jack Phillips), and indeed Kirkpatrick himself, have referred to a possible fifth level, namely ROI (Return On Investment). ROI can easily be included in Kirkpatrick's original fourth level 'Results'. The inclusion and relevance of a fifth level is therefore arguably only relevant if the assessment of Return On Investment might otherwise be ignored or forgotten when referring simply to the 'Results' level.

Learning evaluation is a widely researched area. While Kirkpatrick's model is not the only one of its type, for most industrial and commercial applications it suffices; indeed most organizations would be absolutely thrilled if their training and learning evaluation, and thereby their ongoing people-development, were planned and managed according to Kirkpatrick's model.