Growing Up- Building Learning Performance Measurement Maturity

Peggy Parskey with Kristin Hodges

May 2006

Table of Contents

Table of Figures

Introduction

Within most companies today, senior business leaders increasingly expect the Learning and Development (L&D) function to demonstrate its value to the broader enterprise. The function is under considerable scrutiny. Many corporations are outsourcing portions of their L&D function. Others have reduced their training and development investments even though employee competency improvement is a stated objective of these firms. L&D functions are finding it necessary to show tangible return on the investment in human capital development and simultaneously demonstrate increased operational efficiency and high levels of learner satisfaction with their program offerings.

This climate has made it imperative for the L&D function to measure its impact at a variety of levels: from the most tactical and operational (learning hours delivered), to more management focused (perceived value of the training courses) to the very strategic (business impact). The literature abounds with methods, tools and techniques to apply Kirkpatrick's four levels of training measurement or Phillip's additional level measuring Return on Investment (ROI). Systems and tools have been proliferating and there is no shortage of consultants to help L&D functions chase the holy grail of business impact and ROI measurement.

Some organizations are making progress. The ASTD[1] 2005 State of the Industry Report shows that 23% of solutions from their benchmarked organizations are measuring on-the-job application, up from 15% in 2002. But they also report that for these same firms, only 8% of the programs are being measured for business impact, a mere 2-percentage point increase since 2002. Only 3% of all programs are measuring ROI. What's going on? Why such small gains when the technology, methods and resources are more than available?

The answer seems obvious, but it is rarely discussed in the literature. It is simply that few L&D functions appear to have treated the shift to a measurement culture as a significant change effort. Creating a measurement orientation and culture requires a holistic approach touching all the foundations of the organization, including processes, structures, roles, competencies, systems and tools.
Without a systemic approach to the change, it is difficult to create sustainable improvements in measurement maturity.

---

[1] American Society of Training and Development

This White Paper tells the story of Hewlett-Packard's Learning and Development function to improve its maturity in Learning Performance Measurement through the dedication of a small team over a three-year period.

The state of affairs: 2002-2003

A business in transition

To adequately tell the Hewlett-Packard story, we need to reset the clock to the summer of 2002. The merger with Compaq had just been approved in May and knowledgeable employees were being sequestered in "clean" rooms [2] to determine how the two corporations would merge, redesign or adopt processes, structures and operational approaches for the newly combined company. To integrate new Compaq employees into the workplace, every organization was being restructured and reorganized.

In the midst of this change, the Hewlett-Packard Learning and Development (L&D) organization was in the throes of implementing a new Learning Management System (LMS) that would affect all employees across the company. Along with their colleagues in other functions, the L&D clean room specialists were determining which processes would be adopted and which would be retired.

From a learning measurement and evaluation perspective, the organization was not starting from scratch. In 2000, the Chief Learning Officer (CLO) had created a unified, centrally managed learning organization
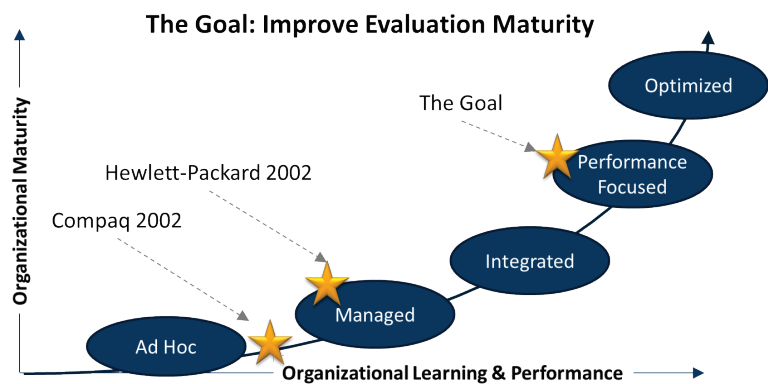
Figure 1: Evaluation Maturity Curve

that consolidated and integrated the disparate training functions that had proliferated across the company. Her goal was to shift the focus from training to performance improvement interventions. To accomplish that goal required consistent processes across the entire learning lifecycle starting with the initial business analysis through

---

[2] The "clean rooms" were segregated working areas for teams designing the operations in the newly merged organization.

design, development, delivery and ultimately support and evaluation. In short, she wanted to increase the maturity of the L&D organization including its maturity with measurement and evaluation. (See Figure 1)

As part of this broader initiative, the CLO had mandated that all training courses use a single evaluation process for measuring student reaction (see Sidebar on page 4) by employing a consistent survey instrument. Despite this corporate mandate, many organizations still used their own processes and instruments, precluding a single corporate view of training effectiveness. Moreover, the newly merged Compaq training organization was still operating with individual training departments, each of which had its own mechanism to survey students and collect reaction data.

In those early days, the Level 1 evaluation processes were entirely paper based. Surveys were distributed at the conclusion of the training to students who filled in the paper forms and turned them into the instructor. The forms were then shipped to a vendor who input the data and provided feedback to the various stakeholders within the organization.

Measuring Training Effectiveness and Impact

Most L&D organizations attempt to measure the effectiveness of their training courses and to use this data to improve their content, structure and delivery. In 1959, Donald Kirkpatrick created a training measurement hierarchy that has persisted to this day. Kirkpatrick defined four levels of measurement:

- **Level 1: Reaction**: This level measures the reaction of students to the course based on a simple survey completed by students at the conclusion of the class. Some L&D practitioners pejoratively refer to these surveys as "smiley sheets".
- **Level 2: Knowledge:** This level measures the increased knowledge that the student acquired during the course. L&D often measures Level 2 by administering a pre and post-test and then measuring the differences between the two results.
- **Level 3: Application:** This level measures the extent to which the student applies the course knowledge back on the job. Level 3 is often measured by surveys, anecdotal feedback or rigorous experimental designs to identify how and to what extent the student has applied the knowledge.
- **Level 4: Business impact:** This level measures the extent to which the training has had a measurable business impact, such as improved productivity, reduced error rates, increased volume of sales or other business metrics
- **Level 5: ROI.** Jack Phillips added this level in the mid-90's to measure the overall return on investment in the training. It uses a technique to isolate the business impacts of the training to compute the ROI on the investment in the design, development and deployment of the program.

Higher levels of evaluation were being employed only sporadically, using no consistent method or instrument. For most L&D professionals, Level 1 reaction data was considered to be sufficient.

However, despite the detailed quarterly analyses and solution level feedback, there was scant evidence that the thousands of surveys collected each month were being

used in a meaningful way. Instructors indicated they reviewed the surveys and made adjustments to their delivery. The most dedicated Solution Owners could attest to program changes that resulted from the detailed review of students' comments. Despite these exemplars, for the most part, L&D was not using the masses of data for making investment decisions or major adjustments to existing programs.

Clearly, if the L&D function were to become a respected organization that produced quality solutions and delivered business value, it had to be more disciplined in measuring the quality of its products. Furthermore, given the huge investment in training across the entire enterprise, the CLO had to utilize higher levels of evaluation to show evidence of business impact to her key stakeholders. Quantitative data from higher levels of evaluation was increasingly required if she was to sustain or grow her L&D budget.

Challenges and opportunities

During the second half of 2002, the L&D function assigned a single manager to be responsible for global training evaluation. The new manager immediately noticed that something was amiss.
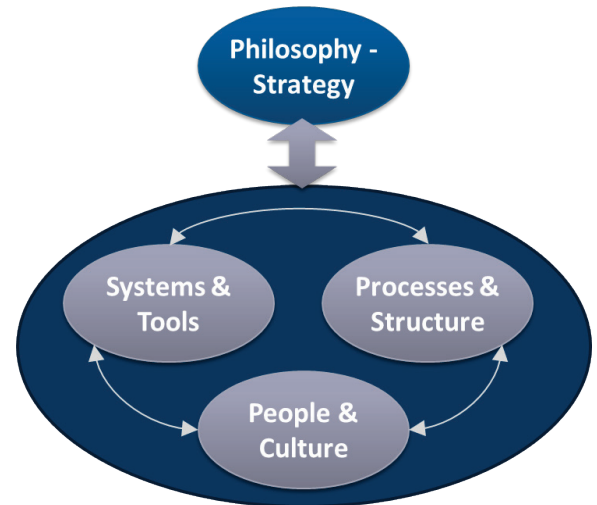
As mentioned earlier, L&D was not sufficiently consistent or disciplined in using its decision-making. Equally important, L&D practitioners viewed evaluation as the last step in the learning lifecycle. For program evaluation to be effective, the owners of the learning solutions (referred to as Solution Owners) had to define success metrics even before a specific intervention was defined. And before launching into the design process, the Solution Owner needed to determine what level of evaluation would be appropriate and how effectiveness would be measured. Without applying rigor at the front end of the process, it would be extremely difficult, if not impossible to conduct a meaningful and credible analysis at the conclusion of the program.

In addition to the lack of consistent processes, instruments and rigor in conducting evaluations, the organization lacked competencies in the process itself. Learning professionals were not schooled in these techniques and many were uncomfortable engaging in a measurement dialogue with their clients in the business. To grow the maturity of the organization would require a major transformation in all aspects of the learning function.

As the CLO and Global Evaluation manager worked together, they began to realize that the challenge before them was multi-pronged. The organization lacked a philosophy about measurement and a corresponding strategy. The CLO believed both were required to transform the organization. Beyond the strategy, the Global Learning

Process team had to redesign processes, create a structure, address competencies and cultural norms, create tools, and job aids and finally, implement enabling systems to simplify the complex approaches to evaluation and measurement across the organization (see Figure 2).

The Team Creates a Strategy

Early in 2003, the CLO and Global Evaluation manager jointly created the measurement philosophy and strategy.

Figure 2: Relationships of People, Process and Technology to Strategy

The philosophy contained three critical components:

- **Shift the focus and ensure that all initiatives start with an integrated evaluation strategy.** For the first time, senior leaders in the organization stressed that measurement would no longer be viewed as the last step in the process.
- **Ensure the business defines the success metrics.** This statement was also a significant shift. While pockets of the organization had close working relationships with their business clients, many did not. This philosophical statement implied a significantly new level of engagement with the business client.
- **Leaders make decisions about solutions and their effectiveness based on sound evaluation data.** This statement was also fairly radical. Data had been collected for years, but it was not being used consistently across the organization or up and down the management chain. The CLO wanted to shift the decision making process away from gut feel to a data driven approach.

The strategy identified five key areas for organizational focus:

- **Implement common governance, methods, tools, systems and processes.** This strategy required that all segments of the combined 800+ person L&D organization adopt common approaches for all evaluation and measurement efforts.
- **Use a Human Performance Improvement approach.** As mentioned earlier, the CLO saw the need to move the mission of her organization beyond delivering training to improving employee performance. This implied that performance gaps should be addressed not simply by training, but by a suite of interventions to resolve the problem and sustain the new behaviors. The evaluation approach

therefore needed to look beyond evaluating the training solutions to evaluating the broader solution suite.

- **Improve L&D competencies in measurement and evaluation.** As the Global Evaluation manager had discovered when she assumed the role in 2002, most L&D professionals had only a basic knowledge of evaluation and needed to become much more skilled in the processes and approaches used industry-wide.
- **Assess solution results quarterly at an enterprise level.** While the CLO had quarterly reports and data, she wanted more robust reporting, action planning and decision making to be derived from this data. Moreover, her goal was to report not just Level 1 reaction data, but higher-level evaluation results as well.
- **Invest in solutions that accelerate performance of our people and the business; drive solution portfolio improvements and vendor selections based on evaluation data.** This strategy was intended to ensure that the data would be used not just for making improvements to the training courses, but equally important, that it would help drive investment decisions for the business.

The philosophy and strategy were widely communicated and shared with key individuals in the L&D business units who were champions and possibly early adopters of a more unified approach. At this point, it was now incumbent on the Global Evaluation Team to implement the strategy and drive change within the organization.

Processes and structures

Designing and implementing new evaluation processes

The first order of business for the new Global Evaluation Manager was to design evaluation processes with three critical characteristics:

- They had to be integrated into the Learning Lifecycle processes
- They had to shift the behavior of L&D professionals to initiate evaluation efforts at the very beginning stages as the solution was being explored
- They had to be consistent with a Human Performance Improvement approach such that they could measure the impacts of the training solution, when appropriate, but also scale to measure the impacts of a broader intervention suite as well.

Analyze	Select	Design	Develop	Implement	Support	Evaluate
Identify success criteria using gap analysis	Define evaluation approach & plan	Design evaluation elements	Develop evaluation instruments	Execute the evaluation plan	Support evaluation processes	Analyze & report evaluation results

Business Initiative

Formative Evaluation

Individual Solution Level

Criteria Testing and Certification

Summative Level 1

Summative Level 2

Program/ Initiative Level

Level 3 data feedback loop — Summative Level 3

Level 4 data feedback loop — Summative Level 4

Figure 3: End-to-end Evaluation Model

The timing for this effort was ideal. The Global Evaluation Manager was also responsible for the deployment of a broader Human Performance Improvement process for the entire Learning Lifecycle. Over the prior year, her team had developed and deployed a web-enabled set of tools, processes, templates and guides to help L&D professionals take a broader business view in their design of performance improvement interventions. Solution Owners and internal Business Consultants were trained and mentored on the new approaches. Integration of the evaluation processes was a natural extension of the initial toolset.

The team depicted the Learning Lifecycle and identified the key process steps required for evaluation within each phase (see Figure 3). They then augmented the existing tools and provided additional templates and guides for the Solution Owners, Business Consultants and other key players in the evaluation process.

This was the first time the organization had deployed a complete set of processes that not only provided a step-by-step approach, but equally important, made it clear that evaluation could not be considered the last step in the process, regardless of the level of evaluation being considered.

Developing a data decision and reporting framework

The next challenge for the team was to develop a framework for reporting. Even though the primary data available was still Level 1 reaction data, the team felt confident that this seemly tactical data could help drive strategic decisions.

The tendency for the organization had been to design reports and then determine who got what report. The team decided to take a different approach. They asked critical questions about each management level in the organization. Specifically, they wanted to know:

- What decisions must be made with this data? What actions are you likely to take?
- Given those decisions and actions, what data is required so you can formulate your decisions and construct your action plans?
- How frequently do you need this data to make meaningful decisions?
- In what form must the data be presented to enable you to easily make the decisions and take further action?

FROM	TO
What reports are available?	What decisions must be made? What actions must be taken?
Who should get them?	What data or information is required to make these decisions and take these actions?
	How frequently must these decisions be made?
How frequently should they be distributed	In what form must the data be presented to enable this decision-making
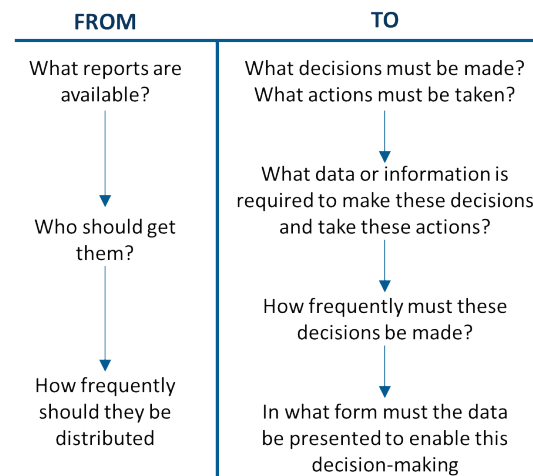
Figure 4: From – To Approach

The mantra for the team became "Start with the end in mind and work backwards from there". They turned the usual reporting discussion upside down by starting with the decision making process first. In addition, the team reinforced the concept that data is only useful if it provides insights into individual, team or organizational performance (see Figure 4).

Creating a virtual organizational structure

The Global Evaluation Team had some inherent challenges based on their position with the organization. They reported into a shared services group that provided a wide range of operational services and capabilities to the organization.

The evaluation initiative required the full engagement of the L&D professionals connected to the business or functional teams. These individuals were advocates for the broader business needs and had a much clearer picture of the challenges and opportunities for greater business alignment and delivering measurement value to their business partners. Without the support of the business-aligned L&D professionals, the long-term efforts would fail.

Later in 2003, the Global Evaluation Manager formed an organization-wide Evaluation Steering Committee. This committee was comprised of evaluation champions within each business unit and the major functional groups (e.g., professional

development, technical training, supply chain). The members were located across the globe and in many cases had not met the other participants. The group met virtually bi-monthly using telephone conferences and virtual classroom facilities to generate and test ideas, and design new approaches that could be deployed within their organizations. Equally important, these individuals were the connection points back into the business and would become champions of the evaluation cultural shift within their organizations. As new systems, processes or tools were deployed, these individuals were also expected to program manage the implementation.

Shifting the culture

Defining organizational accountabilities

As the Global Evaluation Team began addressing the people and cultural aspects of the shift, it became apparent that most L&D professionals were not clear on their accountabilities within the evaluation process, or for that matter, within the broader Learning Lifecycle.

It was critical that the key stakeholders understand what was required of them in each phase of the lifecycle so they could apply these processes in their daily work.
The Global Evaluation team and the Evaluation Steering Committee collaboratively drafted an initial set of accountabilities for the key roles in L&D (see examples in Figure 5). The team then began to present the initial draft with L&D professionals in the functions and businesses.

Evaluation Accountabilities

Task \ Role	Business Consultant	Solution Owner
Gaining Sponsor Agreement	Identify solution success metrics with the project sponsor	
Developing the Evaluation Approach	Develop plan for design/ development of evaluation instruments	Execute the evaluation plan
Implementing the Plan	Ensure data collection mechanisms are in place	Enlist stakeholder support in the implementation of the evaluation plan
Analyzing Results	Analyze evaluation results: define action plans	Analyze evaluation results: define action plans

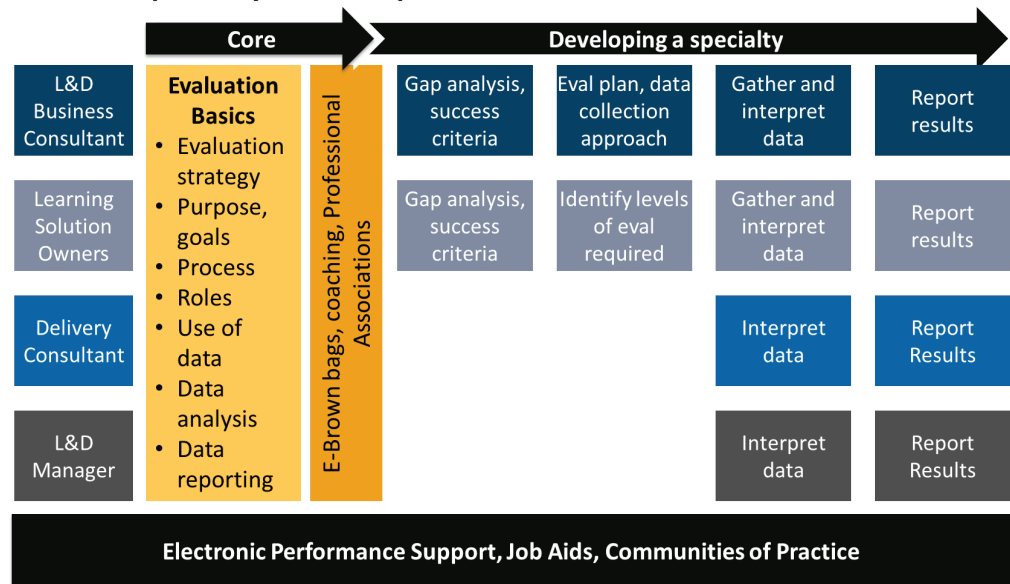*Figure 5: Evaluation Accountabilities*

In most cases, the matrix was an accurate reflection of the assumed (but not explicit) responsibilities of the L&D business teams. To make the accountabilities more meaningful, the business-aligned L&D teams were asked to modify the matrix to reflect both the specialized L&D roles within the business and develop accountabilities that reflected day-to-day work. The goal was not to have a single, consistent set of accountabilities, but rather to ensure each L&D group had documented accountabilities that were appropriate to their needs and aligned with their work processes. By enabling the L&D professionals within the businesses to get "their fingerprints" on the matrix, it was believed that the adoption of these accountabilities would more likely occur.

Building organizational competencies

The Global Evaluation Team was well aware that competencies in measurement and evaluation were insufficient to drive the sustained change required in the organization.

The team included several experts in instructional design and together they developed a roadmap that identified the key roles along with capabilities that each would require to fulfill their accountabilities in evaluation (see Figure 6).

HP's Competency Roadmap

	Core		**Developing a specialty**			
L&D Business Consultant	**Evaluation Basics** • Evaluation strategy • Purpose, goals • Process • Roles • Use of data • Data analysis • Data reporting	E-Brown bags, coaching, Professional Associations	Gap analysis, success criteria	Eval plan, data collection approach	Gather and interpret data	Report results
Learning Solution Owners			Gap analysis, success criteria	Identify levels of eval required	Gather and interpret data	Report results
Delivery Consultant					Interpret data	Report Results
L&D Manager					Interpret data	Report Results
Electronic Performance Support, Job Aids, Communities of Practice						

Figure 6: Measurement & Evaluation Competency Roadmap

The Instructional Designers conducted a series of interviews and concluded that all roles needed an awareness of the L&D measurement strategy and its purpose and goals within the organization. Additionally, all roles would need to learn the basics of evaluation starting with the processes and including data analysis and interpretation.

The team developed and delivered a synchronous e-learning course to L&D professionals across the globe over a 6-week period. The student learning objectives included identifying solution success metrics at the outset, identifying the appropriate level of evaluation, planning the evaluation approach and identifying actions based on evaluation outcomes. Moreover, the course designers directly tied these learning objectives to the evaluation strategy that had been established with the CLO: shift the focus to the beginning of the process, plan the evaluation and then use the data for taking action and making decisions.

The team augmented these courses were augmented with an "e-Brown bag" lecture series (see next section) and offerings from professional associations. The Global Evaluation team also offered mentoring to teams and individuals who needed indepth assistance.

Approximately 100 out of the 800+ organization attended these sessions. The training did not of course create experts overnight. However, the sessions did increase the level of process understanding and triggered many Solution Owners and Business Consultants to try their hand at higher levels of evaluation.

The roadmap also identified more in-depth expertise required of specific roles. For example, Business Consultants, who had the closest relationship with the business client and needs, should be proficient in performance gap analysis techniques and identifying success metrics for the intervention. Managers were expected to interpret data and know how to effectively report and communicate results throughout their own organizations.

The eBrown bag series
As a compliment to the more formal virtual classroom experience, a member of the team spearheaded a series of virtual "eBrown Bag" meetings. They modeled these virtual meetings after the face-to-face lunch time "brown bags" that many HP communities used for presentations or discussions on relevant and leading edge topics.

The team scheduled internal and external speakers to talk about evaluation and its importance not just to the community of Learning Professionals, but to their business partners as well.

The eBrown bag series served a dual purpose. Through the continued presentation of topics on evaluation, the sessions reinforced the importance of measurement skills to the organization at large and the Solution Owners and Business Consultants specifically. In addition, the eBrown bags served to identify the evaluation champions across the geographies and business units. As the intensity of the evaluation efforts increased, the Global Evaluation Team was better positioned to expand their reach through these self-identified supporters.

Implementing a new instrument

A short history of Level 1 surveys

The most basic survey for measuring training satisfaction has been in existence for many years. Dubbed the "Level 1" survey, it is intended to capture student reaction data immediately following a training event (see Sidebar Page 4).

Because these surveys are relatively easy to administer and analyze, nearly all organizations have implemented them to some degree. The 2005 ASTD State of the Industry Report identified that 91% of all firms measure learner reactions after the training event.

At the same time, most organizations consider Level 1 surveys to be marginally useful. Because they do not measure application back on the job or the potential business impact of the training, they are seen as the least important form of training feedback data.

Conducting a statistical analysis

The Hewlett-Packard team took a decidedly different view. They viewed the Level 1 surveys as the employee equivalent to customer satisfaction surveys. These external facing surveys ask many of the same questions (e.g. satisfaction with the product, ease of use, the buying experience) and the data gleaned from customers are treated like gold. Why not use the tens of thousands of surveys in the same manner and gain deeper insights into the views of learners on their training experience?

The first order of business for the team was to ensure that the survey instrument itself produced valid, reliable and unbiased data. The instrument had to be more than a vanity survey to make Solution Owners and leaders look good. They needed to shift the perception that the survey was simply a "smiley sheet" that identified learner happiness with the food, the venue or the instructor. Instead, the survey had to provide meaningful data that would drive decisions and actions.

The existing survey instrument had been tweaked and modified over the years. The services and consulting vendor who had been collecting the surveys, analyzing the data and reporting the results had helped the L&D team make subtle alterations to address the changing needs in the organization. However, in recent months, there was pressure to reduce the number of questions and to make the survey more meaningful. The Global Evaluation Manager and the evaluation consultant/vendor jointly decided that a statistical analysis of the instrument was a key step. This analysis would identify not only how to strengthen the instrument but also how the data could be reported in a meaningful way. The results of the analysis and the subsequent data modeling turned out to be a tipping point in the cultural shift.

The consultant recommended two significant changes:
- Improve the survey by eliminating redundant questions and changing the wording to improve discrimination of learner responses (to reduce positive bias in the responses).
- Improve data reporting by depicting the results using the data model that had emerged from the statistical analysis.

The Global Evaluation Team agreed completely with the consultant's recommendations and took the next steps to implement the revised learner survey and roll out the simplified reporting approach to L&D management.

Creating a data model

Based on the consultant's recommendations, the survey was reduced to a manageable eleven-question instrument. The data model that emerged is depicted in Figure 7. The outside ten boxes and the center box each represent a question on the survey. The analysis revealed that the nine of the ten questions clustered together and the resulting clusters were predictors of overall quality (the center box).

The analysis also revealed that three questions explained 80% of the variability of overall results. These questions were:
- Did this course meet its stated objectives?
- Did this course enhance your job / role performance?
- Would you recommend this course to your peers?

Ironically, these three questions are often the most critical questions on customer satisfaction surveys (i.e. did this product meet your expectations, did it help you do your job and would you recommend it to peers?) Moreover, the former CLO had deemed these same questions as the most important questions to track and she had included them on the organization's balanced scorecard. The "face validity" of the

Level 1 Data Model
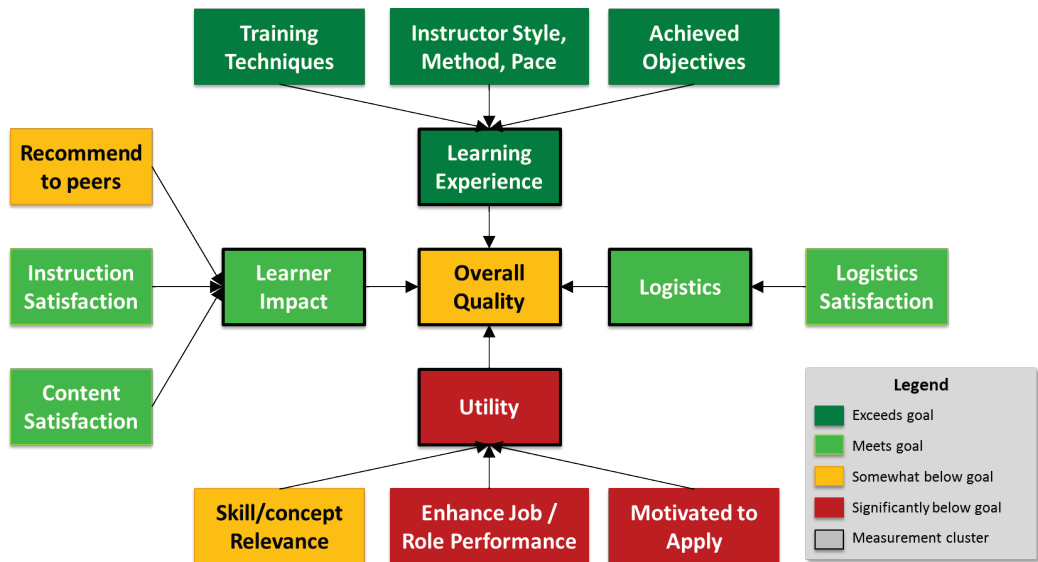
Figure 7: Measurement Data Model

findings and the ease of depicting results became a significant factor in the coming months.

Getting traction

With a simple way to represent results, the Global Evaluation Team published solution and portfolio performance using the data model framework.

Each block in the data model was color coded to represent the level of performance. Blocks were colored green when the component met or exceeded the performance standard. Blocks were colored yellow when the component narrowly missed the performance threshold and thus were areas for concern. Blocks were colored red when they significantly missed the performance threshold and therefore required immediate action.

When the Global Evaluation Team presented the data in this manner to the L&D leaders and their Portfolio Managers, they observed a noticeable and immediate shift in behavior. Now, the leaders could easily gauge their performance against their peers. This view of their relative performance resulted in them paying greater attention to the data and cascading the relevant pages to their staff for explanations and corrective action plans.

For the first time, the Global Evaluation Team presented data in a manner that was simple yet meaningful. This data gave Learning Leaders a high level snapshot and

enabled them to determine where to put their energies and when they should take remedial actions.

The data model representation triggered deeper and more frequent discussion. L&D Leaders regularly invited members of the Global Evaluation Team to staff meetings to discuss the results and formulate hypotheses on why various factors were performing well or not meeting performance standards.

The data model also triggered questions for more detail. Leaders wanted access to the details, including student comments. They also wanted the ability to "double-click" on a particular data point and drill down for additional granularity. Suddenly, data that had been used only sporadically was being examined in great depth, by more individuals and at higher management levels than ever before. Clearly, this approach to data reporting and analysis had created a higher energy level in the organization.

Moving to higher levels

The Evaluation Team needed to capitalize on this newfound energy and interest in evaluation data.
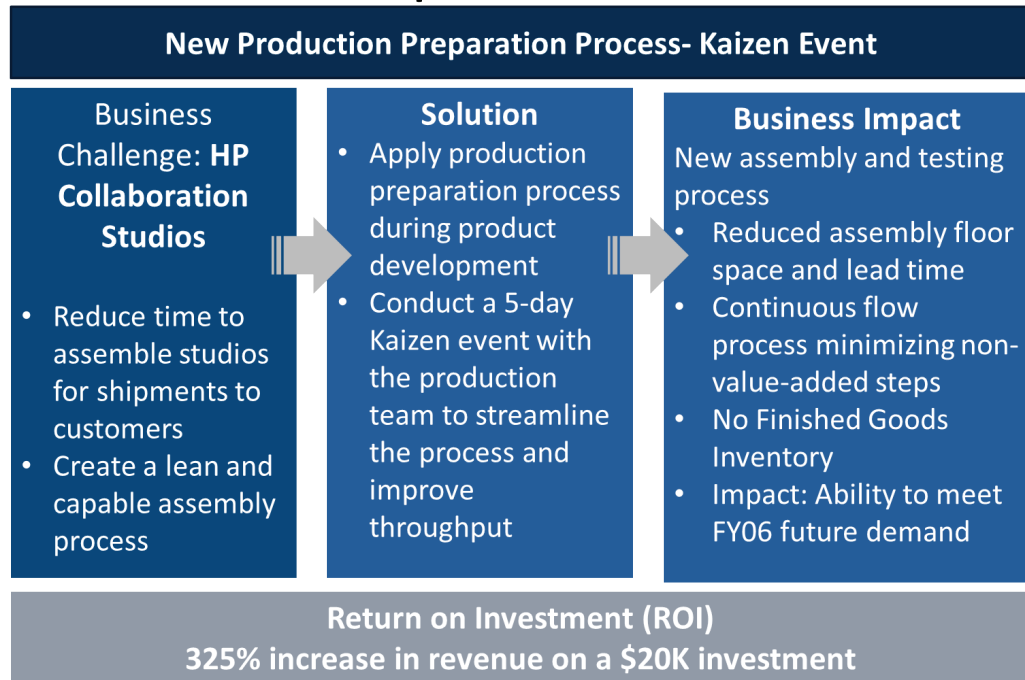
When the CLO left the company, her replacement focused on the numbers and having solid evidence to support anecdotal evidence. She was feeling pressure to quantify the value of her organization and began to ask her staff to show evidence of the business impact and ROI of their most critical solutions.

One of the business units was ready to respond to her requests. As the focus on evaluation heated up, the L&D team within the Imaging and Printing Group (IPG) had begun to collect exactly the type of data that the CLO needed.

Lacking automated tools, they had implemented a manual process to quantify and report the business impact and ROI for strategic and critical training programs. These "proof point" stories used a simple framework to identify the business problem, the solution, the impact of the solution on the business results and finally the return on the training investment (see Figure 8). The L&D senior leaders in the Imaging and Printing Group used these stories in discussions with their business partners to demonstrate the value of the L&D organization to the business. They were also provided to the CLO who began asking the other members of her staff for similar data.

Increasingly, as the demand for business impact and ROI data increased, the Global Evaluation Team was asked to guide Solution Owners and help them develop an evaluation approach and plan that was appropriate to their solution.

ROI Proof Point Example

New Production Preparation Process- Kaizen Event		
Business Challenge: HP Collaboration Studios • Reduce time to assemble studios for shipments to customers • Create a lean and capable assembly process	**Solution** • Apply production preparation process during product development • Conduct a 5-day Kaizen event with the production team to streamline the process and improve throughput	**Business Impact** New assembly and testing process • Reduced assembly floor space and lead time • Continuous flow process minimizing non-value-added steps • No Finished Goods Inventory • Impact: Ability to meet FY06 future demand
Return on Investment (ROI) **325% increase in revenue on a $20K investment**		

Figure 8: ROI Proof Point Example

The culture was beginning to shift and with that the need for more automated and simpler approaches to measuring business impact and ROI. The proof point model was gradually becoming the simplified way to depict the impact of these major initiatives.

Implementing automated systems

The L&D organization implements its first automated evaluation system

Up to this point, the organization had no automated system for collecting Level 1 reaction survey data from students. While the standard paper process had been generally adopted across the organization, it clearly did not meet the new demands for fast feedback turnaround, self-service reporting and data filtering to present varied views of the information.

In 2003, the L&D senior management team funded a project to implement an evaluation system. The application chosen had been available in-house for some time, having been bundled into the purchase of the Learning Management System. While the tool appeared to meet the basic needs for survey collection, further investigation revealed that its data reporting capability did not meet the users' needs and would

inhibit the team's ability to meet the increased demand for reports and actionable information.

The team sketched out the project and felt it should be implemented in phases. Phase I would address automated survey data collection and provide basic reporting capability. More sophisticated reporting would be implemented in Phase II. The team's belief was that a phased approach would deliver benefits to the organization much faster and would enable users to become familiar with the system. Users with greater knowledge of the system would also have better defined requirements that could accelerate progress for the Phase II reporting project.

Unfortunately, the systems implementation was fraught with challenges exacerbated by reduced budgets and reallocation of resources to higher priority efforts. Despite the challenges, Phase I was launched in the spring of 2004. As the team knew all too well, the automated solution provided only the most rudimentary reporting and did not enable instructors, Solution Owners or higher level managers to obtain data that would aid decision-making.

As cost controls tightened, Phase II was cancelled and prevented the Global Evaluation Team from providing a robust reporting tool. They feared that the momentum gained in evaluation would be lost with the limited tool available.

L&D implements a second Evaluation System

During the summer of 2004, the Global Evaluation team was presented with an opportunity. The senior leaders of the three Hewlett-Packard training organizations (that served internal employees, customers and channel partners) launched a major Learning Integration initiative. The goal of this initiative was to eliminate or minimize redundant systems, processes and resources and drive toward a unified and standard approach for all recipients of HP training, regardless of their affiliation or origin. One component of this initiative was to implement a single evaluation platform across all three training organizations and standardize on processes and reaction survey instruments.

The Global Evaluation Team saw this initiative as an opportunity to have a broader impact within HP and at the same time provide the needed reporting capability to their organization. With funds being diverted to this cross-organizational initiative, this newly launched cross-organizational Evaluation initiative was viewed as a way to deliver the Phase II requirements even if it meant scrapping Phase I and essentially starting over. To make certain that these needs would be met, the Global Evaluation Manager volunteered to become the Global Program Manager and coordinate the

selection, implementation and deployment of the systems and processes across all geographies and business units in all three organizations.

Even before the program was formally launched, the Global Program Manager had to identify a single online evaluation software solution that would meet the disparate needs of the three organizations. The three organizations operated like five separate entities[3] each of which had its own processes, tools, systems and approaches for collecting, analyzing and reporting data. The Global Program Manager feared that no single application would meet the needs the five entities. Moreover, she was concerned that the superset of requirements would necessitate a project with significant complexity and a vast scope that would undermine her ability to implement this capability in a short time frame.

As the Program Manager began discussions with her colleagues in the other organizations, she found that a solution[4] was already available[4]. She herself had investigated this solution a year earlier and believed it met the functional requirements of her organization. However, she did not pursue this avenue at the time because her team was already in the throes of implementing the resident automated evaluation system.

She was pleased to find, however, that her European colleagues in the partner organization were already implementing this alternative system in their region. In addition, she learned that the Customer Education Learning Organization was also about to launch an implementation with this same system. She and her extended team decided to conduct a very basic needs analysis and make a decision. Within a week, they decided that they had enough information and concluded that this system would meet the needs of all five entities.

---

[3] The partner organization operated at a regional level, each of which had its own system.
[4] Metrics that Matter from KnowledgeAdvisors

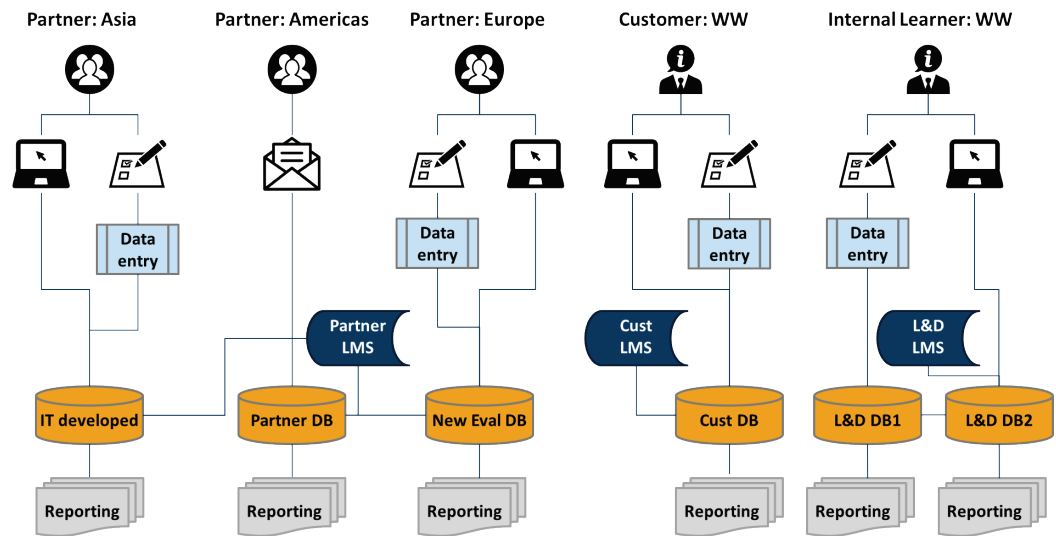## Evaluation System: Prior State

Figure 9: HP Evaluation Systems- Prior State

With the systems decision finalized, the next challenge for the Program manager was identifying how to create a unified view of evaluation and measurement across all five entities (see Figure 9). Some organizations operated globally, others were regionally based. Some had robust reporting and improvement processes, others were less mature. Some organizations had little incentive to move to a new system given their low operating costs or satisfaction with the status quo. Despite the organizational mandate to move forward, there was still some selling to do internally.

The Program Manager identified the key individuals who were least enthusiastic about migrating to a single platform. She met individually with each of the managers to understand their concerns, identify how to mitigate the resistance and get higher-level management support to remove obstacles to success. In each case, she was able to identify an approach that would meet the needs of the broader program while also addressing the concerns of the specific organization or regional team. She was now ready to launch the initiative.

Establishing a cross organizational governance and implementation model

The Program manager formed a team with five project managers responsible for each segment of the business. The collective team was comprised of about 30 individuals including IT representatives, business process designers, application specialists and experts in evaluation distributed across the globe.

In addition, the Evaluation program was aligned with several other initiatives and organizations. In addition to the connection to the larger Learning Integration Program, the Evaluation program was linked to a Joint steering committee, which connected to the stakeholders within each business.

Figure 10 depicts the organization of the project to ensure tight linkages with all the affected stakeholder groups.

Evaluation Governance Model
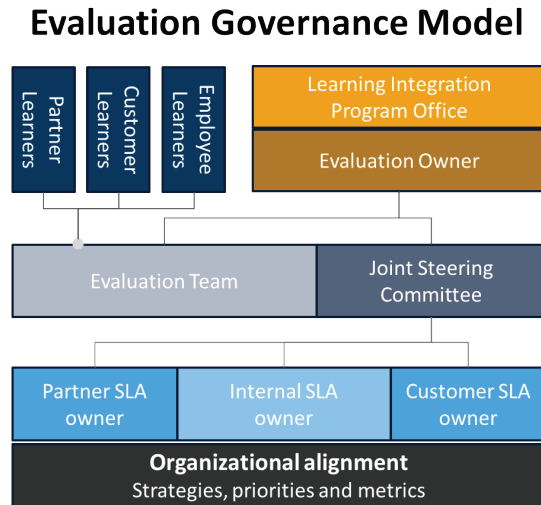
Figure 10: New Evaluation Governance Model

Managing the project

Once the project was formed, the next order of business was to gain agreement by the team on a single survey instrument that would be used across all organizations. As was discussed earlier, evolving the L&D organization to use a single reliable and valid Level 1 reaction survey instrument had taken significant time, energy and resources. The Global Program Manager now had to revisit the entire discussion with the five project teams each with strongly held ideas of their own.

After lots of negotiation, compromises and input from an expert in evaluation instrument design, the team finally agreed on the core surveys that would be used consistently with all three organizations across the globe. With this linchpin in place, the team was now ready to move ahead with the technical details of the implementation, which included translating the instrument and user interfaces into 20 separate languages that were most prevalent by the learner community.

The schedule was tight and required that the largest implementation, for the HP employee population of 150,000 be deployed in less than six months. The other implementations were staggered to enable best practices from earlier deployments to be applied to subsequent implementations and to ensure that vendor resources could be allocated to assist with the rollout.

Although the project had its challenges and unforeseen obstacles, it was completed in the expected timeframe, within scope and with the planned budget. The team had successfully met the challenge of implementing a single IT system for Learning and Development that crossed all organizational boundaries.

Measuring success

Program results

A program focused on improving measurement maturity should of course, measure its own success and organizational impact. The results are summarized in Figure 11.

The Program Team was proud to report that the results were not only quantifiable savings in operational costs but also external recognition by the leading organization in the field of training and development (see below).

Result	Metric
Eliminated redundant processes and systems	$1M over 5 years
Eliminated paper processing and vendor costs	$1M over 5 years
Improved perceptions on utility for e-learning courses	20 percentage points over 18 months
Increased volume of reliable and valid data	80K evaluations/quarter 4-fold increase in 18 months
Improved organizational maturity in evaluation	Named by ASTD as #2 BEST Award winner in 2005
Client discussions focused on business issues	Priceless

Figure 11: Tangible and Intangible Benefits

Moreover, the largest and most sustainable impact was the shift in the dialogue between the L&D professionals and their business owners. The conversations significantly moved from promotion and defense of L&D value to a focus on the impact to the business, collaboratively defined success metrics, and reporting in the language of the business client. As the Evaluation Program Manager said, the value of this shift was (to borrow from the ubiquitous Master Charge commercial), "priceless".

The maturity journey

By the end of 2005, HP's Learning and Development organization had made substantial progress. The organization was now measuring its programs with greater regularity. As a testament to this progress, the American Society of Training and Development (ASTD), named Hewlett-Packard a BEST[5] Award winner in 2005,

---

[5] **The ASTD BEST Awards** recognize organizations that demonstrate enterprise-wide success impacted by employee learning and development.

honoring the L&D firm with the second spot on the list, up from 14th place the prior year.

L&D leaders agreed that multiple measures demonstrated that the organization had moved up the maturity curve, and was now much closer to its goal.

Critical success factors

As the team reflected back on the critical success factors for this program, they cited several key factors that enabled the program to achieve its goals:

The Journey Up the Measurement Maturity Curve

Figure 12: The Measurement Maturity Journey

- Sponsorship: The sponsorship from the CLOs and senior leaders made it possible to sustain momentum, secure the needed funding to implement new processes and tools and ultimately, reinforce the importance of measurement to the organization.
- Inclusiveness: Champions from all parts of the organization were invited to participate. The inclusive approach embraced all business units, geographies, views and roles.
- Pull vs. push: The team instinctively knew that a push approach to implementing change could not be sustained. At each point along the way, they sought out local champions with tight connections to the business to create the pull for the processes, data and tools.
- Build a credible solution: While the evaluation team had significant experience in the field, they also realized that their knowledge was limited and biased by their prior experiences. The team continually engaged external experts to validate their work, provide new insights and keep the team from "breathing their own exhaust" as one team member noted.
- Engage resisters: As one wise person once said, "You know you are getting close to the target when the opposition starts firing back". Rather than view resistance as a negative reaction, the team chose to embrace the resistors, understand their issues, keep them involved and eventually turn them into supporters.
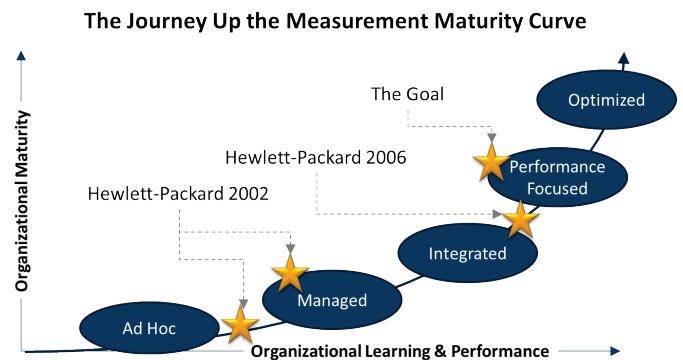
What's next?

Level 1 evaluation:

As with any journey, work continues at Hewlett-Packard. The team continues to improve Level 1 reporting and access to information. They are developing processes to provide role-based reports so that the information presented represents what is most important to each stakeholder group. Moreover, they are continuing to investigate ways to leverage best practices across the organization and engage a community of evaluation champions.

At the same time, the team is addressing the inevitable challenges that arise when implementing major change initiatives. They are continuing to educate the organization and get everyone to embrace the self-service reporting model. And they are also surfacing areas of resistance and engaging those individuals and groups in the planning and ongoing improvement of the processes, tools and systems.

Higher levels of evaluation

For higher levels of evaluation, the team continues to move aggressively. They are currently conducting a pilot with the new tool to collect business impact and ROI data from employees and their managers 3 months post program. This data will enable the organization to collect data systematically and consistently across all critical programs. Moreover, the Evaluation Team will now have the ability to aggregate this data at higher levels and provide a high-level snapshot of impact results to senior management.

The team is also improving the higher-level evaluation tools and templates created in the early stages of the program and making the entire process easier to understand and apply. The team is to developing and provisioning a suite of higher-level evaluation methods and tools that can be selected and applied appropriate to the situation, the business needs and the type of solution being evaluated. In some cases, the automated system will provide the needed capability, but in others, particularly if a robust study is required, the suite will provide the methods and guidelines for conducting a higher-level evaluation study.

Finally, the team will also be looking at the relationship between the Level 1 "utility" data and the Level 3 "application on the job" data to determine to what degree the Level 1 information can provide a low cost early warning indicator to Solution Owners and Portfolio Managers.

Final thoughts

Upon reflection about the HP L&D measurement transformation, three aspects should be noted:

First, a sustainable measurement program must take a broad organization design approach. By addressing all the organizational drivers (e.g. processes, structure, roles, competencies, sponsorship) and aligning them with the strategy, measurement can become part of the fabric of the organization. Addressing only a few of these components can produce results, but they are often short lived and viewed as the "flavor of the month".

Second, it is about managing change. This includes the need for senior sponsorship and establishing clear governance. Moreover, any change initiative must embrace cultural perspectives and build trust throughout the team, be they virtual or co-located. And as has been noted, successful change projects build solid allies and engage the resistors. For this project in particular, there were so many changes underway that this aspect had to take priority. The merger cost cutting, reallocation of resources, systems integration and shifting priorities made it even that more critical that this effort be handled like a major change initiative.

Finally, it is about program and project management. The program needs to be clear about success metrics and be explicit about what "done" looks like. Project managers need to manage assiduously the scope, schedule and resources as the organization changes, new requirements emerge or unexpected issues surface. Like any large project, the plan should be phased to produce small wins along the way and demonstrate tangible progress to sponsors and stakeholders. Finally, project roles, responsibilities and accountabilities must be made explicit, particularly when teams are virtual and do not report to the Program manager.

Managing these three dimensions made it possible to drive this change over a three-year period and make strides that few would have thought possible at the outset. The team is proud of what they accomplished and look forward to helping the organization continue its journey up the maturity curve.

Contact Me

✉ Email: pparskey@mjpnetwork.com

🌐 Website: www.parskeyconsulting.com

in LinkedIn: https://www.linkedin.com/in/peggy-parskey-11634

🐦 Twitter: @peggyparskey