Measuring eLearning to Create Virtuous Cycles of Continuous Improvement

Will Thalheimer, PhD
President
Work-Learning Research, Inc.
Somerville, Massachusetts, US

Phone: 888-579-9814
Email: info@work-learning.com
Website: Work-Learning.com
Audits: LearningAudit.com
Smiles: SmileSheets.com
Blog: WillAtWorkLearning.com
Twitter: @WillWorkLearn

Bridging Gap between Research and Practice
Slides available at: www.is.gd/willstuff
Simplicity is the Ultimate Sophistication

Leonardo da Vinci

The Decisive Dozen
for Learning Design and Learning Measurement

1. Content
2. Exposure
3. Guiding Attention
4. Creating Correct Conceptions
5. Repetition
6. Feedback
7. Variation
8. Retrieval Practice
9. Context Alignment
10. Spacing
11. Persuasion
12. Perseverance

http://is.gd/DecisiveDozen
http://is.gd/ddResearch
The 22 Principles

Available at: is.gd/manifesto22
Quite simply, the BEST book on smile sheet creation and utilization, Period!

Karl M. Kapp
Professor of Instructional Technology
Bloomsburg University

Thoughtful and sensible advice for feedback tools that will provide valid and actionable data.

Robert O. Brinkerhoff
Professor Emeritus, Western Michigan University & Director, Brinkerhoff Evaluation Institute

Evidence-based practice at the master level.

Julie Dirksen
Author of Design For How People Learn
How effective are your organization’s smile sheets—those that you use in your elearning?

1. We DO NOT USE smile sheets in our elearning.
2. We DO NOT USE THE DATA we collect.
3. Our smile sheets are NOT VERY EFFECTIVE
4. Our smile sheets are VERY EFFECTIVE

How do you typically measure Level 2 Learning Results—in your elearning?

1. With items that focus on KEY TERMINOLOGY
2. With items that focus on KEY CONCEPTS
3. With REALISTIC SCENARIO-BASED DECISIONS
4. With SIMULATIONS or HANDS-ON EXERCISES
Which of the following practices is done routinely in your organization to measure learning results?

1. We measure learning results a week or more after learning events.
2. We measure post-training on-the-job application of what was learned.
3. We watch learners as they use our elearning programs.
4. We track on-the-job learning.

The Kirkpatrick Model Rules!

Level 1 Reaction
Level 2 Learning
Level 3 Behavior
Level 4 Results
Later on the job, learners remember what they've learned. The learners get a return on their efforts.

Our learners build understanding. Performance can be prompted through job aids, signage, intuitive cues, performance support, management, etc.

Learners can learn on-the-job through trial & practice, insight learning, help from others, social media, studying on their own, etc.

The organization gets a return on its investment. Most Smile Sheets Not Correlated with Learning!
Correlation between levels?

Very Weak Relationship between Levels

Alliger, Tannenbaum, Bennett, Traver, & Shotland (1997).
A meta-analysis of the relations among training criteria.
Personnel Psychology, 50, 341-357.

Correlation between levels?

Weak Relationship is below .30 and .09 is VERY WEAK
So...SMILE SHEETS tell us VERY LITTLE about Learning

The Critical Importance of Remembering
Learning and Forgetting Curves

If our learners start here.

But end up here.

Have we maximized the learning benefits?

There are many possible after-training results, depending on:

(1) Design of the Learning
(2) After-Learning Follow-up

YouTube: http://is.gd/LearningForgettingCurves
What does an end-of-course assessment tell us?
Three Biases in the Way We Measure Level 2 Learning

WHEN did you Measure Learning?

6) If you said you measured Learning in Question 4, WHEN were the evaluation instruments presented to the learners? Check all that apply.

- At the end of the learning event 90%
- Later the same day
- On the NEXT DAY after the learning event ended
- From 2 to 6 DAYS later
- From 1 to 3 WEEKS later
- From 1 to 3 MONTHS later
- From 4 MONTHS TO 1 YEAR later
- MORE THAN 1 YEAR later

Source: The eLearning Guild Research
“Level C represents the last level of certification that can be considered to assess an ability to perform on the job. Level D represents the first quantum jump away from fidelity in assessment and should be used with caution.”

Alena wants to start a firm that helps farmers grow food organically. She has a degree in sustainable agriculture and has worked for a non-profit organization for seven years doing similar work. She’s developed a marketing plan, a financial plan, and has found several farmers who would pay her if she went out on her own. What should Alena do first—before she tells her boss that she’s quitting to start her own firm?

A. Create a cash flow statement to determine whether her predicted income will support the business through the first year.

B. Form a group of advisors with experience in both small-business management and agriculture.

C. Determine whether she has enough seed money and start-up capital to get started.

D. Analyze her values and goals to ensure that the proposed business will support them.

Scenario-Based Question Example
Smile Sheets
Can be Improved!

Likert-like Scales provide Poor Data

Sharon Shrock and Bill Coscarelli, authors of the classic text, now in its third edition, *Criterion-Referenced Test Development*, offer the following wisdom:

On using Likert-type Descriptive Scales (of the kind that uses response words such as “Agree,” “Strongly Agree,” etc.):

“...the resulting scale is deficient in that the [response words] are open to many interpretations.” (p. 188)

**Likert-like Scales provide Poor Data**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td>B.</td>
<td>C.</td>
<td>D.</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>Agree</td>
<td>Neither Agree Nor Disagree</td>
<td>Disagree</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Compared to:
- Previous
- Standard
- Others
We’d like to trust our learners…

But the research shows that they don’t always know their own learning…

**Learners are Overly Optimistic**
Zechmeister & Shaughnessy (1980).

**Learners Fail to Properly Use Examples**
Renkl (1997).

**Learners can’t always Overcome Faulty Prior Knowledge**
Kendeou & van den Broek (2005).

**Learners Fail to Give Themselves Retrieval Practice**
Karpicke, Butler, & Roediger (2009).

**Two Recent Reviews Emphasize Learners’ Lack of Knowledge of Learning**
Transmogrify

We Start with Fuzzy Adjectives

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neither Agree Nor Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Magically We Turn Adjectives Into Numbers

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neither Agree Nor Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0</td>
<td>4.2</td>
<td>4.1</td>
<td>3.9</td>
<td>4.1</td>
</tr>
<tr>
<td>4.1</td>
<td>3.9</td>
<td>4.2</td>
<td>3.7</td>
<td>4.0</td>
</tr>
<tr>
<td>3.4</td>
<td>4.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We Average Responses, Losing More Info

We Choose One Question and Report Results

My Journey in trying To create a better Smile Sheet
My "New" Smile Sheet

Specific Concepts

How much Value?

Concept New?

How Well Taught?

<table>
<thead>
<tr>
<th>Learning Concept</th>
<th>Value of Specific Information</th>
<th>Circle One Below</th>
<th>How new was concept to you?</th>
<th>How well was concept taught?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>How new was concept to you?</td>
<td>How well was concept taught?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>How new was concept to you?</td>
<td>How well was concept taught?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>How new was concept to you?</td>
<td>How well was concept taught?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>How new was concept to you?</td>
<td>How well was concept taught?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>How new was concept to you?</td>
<td>How well was concept taught?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>How new was concept to you?</td>
<td>How well was concept taught?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
<td>Circle ONE Below</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>How new was concept to you?</td>
<td>How well was concept taught?</td>
</tr>
</tbody>
</table>

A. Two common practices in measurement bias our results. Measuring learning at end of program may bias the results. Similarly, measuring learning in the learning context.

B. Measuring retrieval is essential (even if we measure on-the-job performance and results) because retrieval is required for on-the-job application. It is on the causal pathway from learning to performance and results.
Simple Overall Ratings

Helping Learners Calibrate

<table>
<thead>
<tr>
<th>Overall Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>Very Little Value</td>
</tr>
</tbody>
</table>

Rate the overall value of the learning experience. Circle ONE number (Please Don’t Circle the Words). 

Rate your physical comfort during the learning experience. (consider breaks, food, temperature, furniture, lighting, etc.). Circle ONE number.

Likelihood that you will utilize what you learned in the next two weeks. Circle ONE of the percentages.

Likelihood that you will share what you’ve learned with a colleague or friend in the next two weeks.

Best Feedback Comes from Comments

Overall Comments – Please write your feedback here. Very valuable to us!!!

Thanks for your Participation, your Feedback, and your Ideas! Keep in touch (www.willatworklearning.com)
Performance-Focused
Smile Sheet
Quite simply, the BEST book on smile sheet creation and utilization, Period!

Karl M. Kapp  
Professor of Instructional Technology  
Bloomsburg University

Thoughtful and sensible advice for feedback tools that will provide valid and actionable data.

Robert O. Brinkerhoff  
Professor Emeritus, Western Michigan University & Director, Brinkerhoff Evaluation Institute

Evidence-based practice at the master level.

Julie Dirksen  
Author of Design For How People Learn

<table>
<thead>
<tr>
<th>Ultimate Goal</th>
<th>Primary Goals</th>
<th>Secondary Goals</th>
<th>Tertiary Goals</th>
</tr>
</thead>
</table>
| MAXIMALLY EFFECTIVE SMILE SHEET | LEARNING EFFECTIVE?  
Will the learning be effective in supporting on-the-job performance? | RESULTS ACTIONABLE?  
Will the Smile-Sheet results communicate with clarity and urgency to guide action? |

Ensuring that we are focused on the science-of-learning factors that matter!

Three Key Goals:

1. **Are the Questions Well-Designed?**
   - Learners can answer the questions, are the questions focused on most important factors, are the answers calibrated to provide granularity, are leading questions avoided, do questions avoid areas of bias?
   - From the information, can we determine whether a course needs to be maintained, improved, or removed? Are we avoiding numeric averages that discourage a standards-based decision on success and failure?

2. **Are Learners Making Good Smile Sheet Decisions?**
   - Are we using smile sheets to inform our decisions? Should we be using smile sheets to inform our decisions?

3. **Are the Data we’re Getting Clear and Actionable?**
   - Do smile sheet results distinguish between different levels of success? Are we measuring the things that matter? Are we using the Smile Sheet Opportunity to Educate Our Stakeholders?
A Better Smile Sheet Question

In regard to the course topics taught, HOW ABLE ARE YOU to put what you’ve learned into practice on the job?

A. I’m NOT AT ALL ABLE to put the concepts into practice.

B. I have GENERAL AWARENESS of the concepts taught, but I will need more training/practice/guidance/experience TO DO ACTUAL JOB TASKS using the concepts taught.

C. I am ABLE TO WORK ON ACTUAL JOB TASKS, but I’LL NEED MORE HANDS-ON EXPERIENCE to be fully competent in using the concepts taught.

D. I am ABLE TO PERFORM ACTUAL JOB TASKS at a FULLY-COMPETENT LEVEL in using the concepts taught.

E. I am ABLE TO PERFORM ACTUAL JOB TASKS at an EXPERT LEVEL in using the concepts taught.
In regard to the concepts taught in the course, how motivated WILL YOU BE to UTILIZE these skills in your work?

A. I will NOT MAKE THIS A PRIORITY when I get back to my day-to-day job.
B. I will make this a PRIORITY—but a LOW PRIORITY when I get back to my day-to-day job.
C. I will make this a MODERATE PRIORITY when I get back to my day-to-day job.
D. I will make this a HIGH PRIORITY when I get back to my day-to-day job.
E. I will make this one of my HIGHEST PRIORITIES when I get back to my day-to-day job.
Now that you’ve taken the course, how well do you feel you understand the concepts taught in the course?

A. I have some significant CONFUSIONS AND/OR BLIND SPOTS.  

B. I have a BASIC FAMILIARITY with the concepts.  

C. I have a SOLID UNDERSTANDING of the concepts.  

D. I have a COMPREHENSIVE UNDERSTANDING of the concepts.  

E. I have an EXPERT-LEVEL UNDERSTANDING of the concepts.
This learning event will make a significant contribution to my work, in terms of the following: SELECT AS MANY AS YOU LIKE!

A. Improving my personal productivity.
B. Increasing my ability to innovate.
C. Enabling me to collaborate more effectively.
D. Improving my management performance.
E. Enabling me to generate more revenue.
F. Enabling me to lower costs.

The Future of Learning Measurement
Why 'Accelerating Time-to-Proficiency' metrics are important in today's business world?

Nano-Coaching Cycle

- Goal Achieved?
- Perform Task
- Submit Work Product
- Notify Coach
- Coaching Support
  - Checklists, Guidelines
  - Coach the Coach

Specific Feedback
- Photo
- Checklist
- Text
- Comment
- Audio
- Document
- Direct Observation

Direct Observation
- Email & Dashboard

©2014-2015 Cognitive Advisors LLC
Used With Permission

Marty Rosenheck
CEO
Any Learning Experience  xAPI  Learning Record Store (LRS)  Analytics and Reporting

Activity Stream

<Actor>  <Verb>  <Object>
Learner  Played  Simulation
Learner  Cleaned  Crankshaft
+ Context, Results, and Extensions

Mike Hruska
President/CEO

©2015 Problem Solutions Used With Permission

xAPI Example

```
{  
"id": "5d23f3c4-9365-46ad-8390-5a096f5a529",  
"actor": {  
"objectType": "Agent",  
"name": "Smith, Greg",  
"mbox": "mailto:admin@sp2.com"  
},  
"verb": {  
"display": {  
"en-US": "assessed"  
},  
"id": "http://www.SP2.com/assessed"  
},  
"object": {  
"objectType": "Agent",  
"name": "John Bates",  
"mbox": "mailto:john.bates@us.army.mil"  
}
```

www.pipelinexapi.com

Mike Hruska
President/CEO

©2015 Problem Solutions Used With Permission
Human Performance Measurement Language (HPML)

- HPML is an XML Schema-based language.
- “Takes raw data as input and specifies the computations required to turn those data into measurements and assessments.”
- Provides a list of constructs, which can represent experiences, tasks, conditions, etc.
- Based as an XML Schema
- Study underway at Simulation Interoperability Standards Organization (SISO) for standard

References:

©2015 Problem Solutions
Used With Permission
Comparing two versions of a program.

Subscription Learning

- Learners Subscribe or are Subscribed
- Many Learning Events
- Usually Short Nuggets
- Spaced Over Time
- Usually Relies on Push Technology
- Usually Utilizes the Spacing Effect

To Learn More: SubscriptionLearning.com
Subscription Learning can be – THEORETICALLY – a series of informational nuggets

<table>
<thead>
<tr>
<th>October</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuggets</td>
<td>Nuggets</td>
<td>Nuggets</td>
</tr>
<tr>
<td>Nuggets</td>
<td>Nuggets</td>
<td>Nuggets</td>
</tr>
<tr>
<td>Nuggets</td>
<td>Nuggets</td>
<td>Nuggets</td>
</tr>
</tbody>
</table>

Subscription Learning IS MORE THAN INFORMATIONAL NUGGETS

<table>
<thead>
<tr>
<th>October</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Welcome! Challenge Q1 Challenge Q2 Feedback</td>
<td>Challenge Q4 Challenge Q5 Feedback Assignment</td>
<td>CEO Video Challenge Q3 Feedback Reflection</td>
</tr>
<tr>
<td>Challenge Q6 Challenge Q7 Feedback Trainer Video</td>
<td>Assign. Review Challenge Q9 Feedback Assignment</td>
<td>Assign. Review Self-Assessment</td>
</tr>
<tr>
<td></td>
<td>CEO Video Reflection Challenge Q8 Assignment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Challenge Q10 Challenge Q11 Feedback Reinforcer 2</td>
</tr>
<tr>
<td>Challenge Q12 Challenge Q13 Feedback Reinforcer 3</td>
<td>Challenge Q14 Feedback Reinforcer 4 Change Effort</td>
<td></td>
</tr>
<tr>
<td>October</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>Welcome!</td>
<td>Challenge Q1</td>
<td>Challenge Q2</td>
</tr>
<tr>
<td></td>
<td>Feedback</td>
<td>Feedback</td>
</tr>
<tr>
<td>CEO Video</td>
<td>Challenge Q3</td>
<td>Feedback</td>
</tr>
<tr>
<td></td>
<td>Reflection</td>
<td>Reflection</td>
</tr>
<tr>
<td>Challenge Q4</td>
<td>CEO Video</td>
<td>Challenge Q8</td>
</tr>
<tr>
<td>Feedback</td>
<td>Reflection</td>
<td>Assignment</td>
</tr>
<tr>
<td>Assignment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Challenge Q6</td>
<td>Assign Review</td>
<td>Challenge Q7</td>
</tr>
<tr>
<td>Feedback</td>
<td></td>
<td>Assignment</td>
</tr>
<tr>
<td>Trainer Video</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assign Review</td>
<td>Challenge Q9</td>
<td>Feedback</td>
</tr>
<tr>
<td>Challenge Q10</td>
<td>CEO Video</td>
<td>Challenge Q11</td>
</tr>
<tr>
<td>Feedback</td>
<td>Reflection</td>
<td>Feedback</td>
</tr>
<tr>
<td>Reinfacer 1</td>
<td></td>
<td>Reinfacer 2</td>
</tr>
<tr>
<td>Challenge Q12</td>
<td>Challenge Q13</td>
<td>Feedback</td>
</tr>
<tr>
<td>Feedback</td>
<td></td>
<td>Reinfacer 3</td>
</tr>
<tr>
<td>Reinfacer 4</td>
<td>Challenge Q14</td>
<td>Feedback</td>
</tr>
<tr>
<td>Change Effort</td>
<td></td>
<td>Reinfacer 4</td>
</tr>
</tbody>
</table>

Guerrilla Evaluation

- Our goal: Creating better feedback loops.
- Gather satisfactory data at a good cost point.
- Consider taking a user-testing approach.
  - With Online Meeting Tools, it’s easy, it’s almost FREE, it takes hardly any time at all.
- Sample across cohort of users.
- Moral licensing.

How many of you have EVER seen a learner using the elearning you’ve developed?

Julie Dirksen, Author
Design for How People Learn
Problems with the Kirkpatrick/Phillips 4- or 5-Level Models of Learning Evaluation

- Pushes us to focus on weighing outcomes. Is largely silent on learning support and learning-design improvement.
- Training centric. Ignores prompting mechanisms & on-the-job learning.
- Ignores the role that management and the business side must play.
- Implies that higher levels are more important than lower levels.
- Ignores the causal chain from learning to remembering to performance to results.
- Ignores the fact that learners forget and that learning interventions can be good at creating understanding but poor at minimizing forgetting.
- Pushes us to value learner ratings as predictive of learning and on-the-job performance.
“Historically, organizations and training researchers have relied on Kirkpatrick’s [4-Level] hierarchy as a framework for evaluating training programs…

[Unfortunately,] The Kirkpatrick framework has a number of theoretical and practical shortcomings.

[It] is antithetical to nearly 40 years of research on human learning, leads to a checklist approach to evaluation (e.g., ’we are measuring Levels 1 and 2, so we need to measure Level 3’), and, by ignoring the actual purpose for evaluation, risks providing no information of value to stakeholders… (p. 91)

http://is.gd/TrainingResearch2012

“The goal of training evaluation is not to prove the value of training; the goal of evaluation is to improve the value of training.” (p. 94-95)

Tim Mooney and Rob Brinkerhoff
Courageous Training: Bold Actions for Business Results.
Final Thoughts:

We have a responsibility to build learning programs that are effective.

Valid feedback enables improvement.

We must work to get good feedback.

We should aim to create virtuous cycles of continuous improvement.