Performance-Focused Smile Sheets: A Radical Rethinking of a Dangerous Art Form

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
Blog: WillAtWorkLearning.com
Twitter: @WillWorkLearn

Training Conference – February 2015

Performance-Focused Smile Sheets: A Radical Rethinking of a Dangerous Art Form

1 Learning Measurement Background
2 Learning Measurement Failures
3 Performance-Focused Smile Sheets
4 Stealth Messaging

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
Blog: WillAtWorkLearning.com
Twitter: @WillWorkLearn
Let me introduce myself…

Bridge

Research & Practice

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
How would you rate your organization’s smile sheets?

There is NO perfect measurement...

1. Terrible
2. Bad
3. Okay
4. Good
Later on the job, learners remember what they've learned.

The learners get a return on their efforts.

The organization gets a return on its investment.

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

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

Percent of Companies Using “To Any Extent”

1. Classroom – Reaction 81%
2. Technology – Reaction 52%
3. Technology – Learning 43%
4. Classroom – Learning 50%
5. Behavior 15%
6. Results 14%
7. ROI 6%
# 1

Your Smile Sheets Stink!
The Kirkpatrick 4-Level Model

Level 1: Reaction
Level 2: Learning
Level 3: Behavior
Level 4: Results

Correlation between levels?

Level 1 to Level 2: $r = 0.09$
Level 1 to Level 3: $r = 0.16$

Very Weak Relationship between Levels

Correlation between levels?

$r = .09$

Practical Significance

No

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


Measurement Scorecard

<table>
<thead>
<tr>
<th>Level</th>
<th>Scorecard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>66%</td>
</tr>
<tr>
<td>Level 2</td>
<td>46</td>
</tr>
<tr>
<td>Level 3</td>
<td>20</td>
</tr>
<tr>
<td>Level 4</td>
<td>12</td>
</tr>
<tr>
<td>Level 5</td>
<td>6</td>
</tr>
</tbody>
</table>

Not Causal

Don’t Do Much
Likert-like Scales & Numeric Responses create 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**

A. Strongly Agree 5
B. Agree 4
C. Neither Agree Nor Disagree 3
D. Disagree 2
E. Strongly Disagree 1

**Compared to:**
- Previous
- Standard
- Others
We Start with Fuzzy Adjectives

Magically We Turn Adjectives Into Numbers

We Average Responses, Losing More Info

We Choose One Question and Report Results

<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>
<tr>
<td>4.0</td>
<td>4.1</td>
<td>3.9</td>
<td>3.7</td>
<td>4.0</td>
</tr>
</tbody>
</table>

# 3

Don’t Forget Remembering!
The Learning Landscape

Learning Intervention → Remembering → On-the-Job Performance → Learning Outcomes

Performance Situation

Individual Results
Organizational Results

Learning and Forgetting Curves

If our learners start here. But end up here. Have we maximized the learning benefits?

Learning and Forgetting Curves

There are many possible after-training results, depending on:
(1) Design of the Learning
(2) After-Learning Follow-up

Learning Curve
On-the-Job Learning Curves
On-the-Job Forgetting Curves

During Learning After Learning
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
0 20 40 60 80 100

# 4
Level 2 Learning Measures are Biased.
Which learning intervention creates the best learning results?

A. Learning Intervention A  
B. Learning Intervention B  
C. Some other answer…(?)

Time Sequence of How Learning Becomes Performance

1st Event

Learning

2nd Event

Retention

3rd Event

Memory Retrieval

2 weeks
Time Sequence of How Learning Becomes Performance

1st Event
Learning

2nd Event
Retention

3rd Event
Performance

What does an end-of-course assessment tell us about what will happen to A and B?

Learning and Forgetting Curves

Biased Metric!!

Understanding

Remembering

Application

During Learning

After Learning

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
- 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
WHERE did you Measure Learning?

7) If you said you measured Learning in Question 4, HOW SIMILAR was the ASSESSMENT CONTEXT to the LEARNING CONTEXT?

<table>
<thead>
<tr>
<th>Context Similarity</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mostly in the Same Context</td>
<td>53.4%</td>
</tr>
<tr>
<td>Mostly in Similar Context</td>
<td>37.1%</td>
</tr>
<tr>
<td>Mostly in Different Context</td>
<td>6.8%</td>
</tr>
<tr>
<td>Not Applicable</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

91%

Source: The eLearning Guild Research

Tested in the learning room or in a different room. Does this matter?

“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.”

# 5

You’re Getting Poor Feedback!
Measurement Scorecard

Level 1  66%
Level 2  46%
Level 3  20%
Level 4  12%
Level 5  6%

Not Causal
1. Biased Timing
2. Biased Context
3. Poor Questions
Don’t Do Much

Are we getting good feedback?
FAIL
### The Instructional-Design Process

<table>
<thead>
<tr>
<th>Common-Sense-Based Instructional Design</th>
<th>Knowledge, Wisdom, Biases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience-Based Instructional Design</td>
<td>Prepare, Analyze, Set Goals</td>
</tr>
<tr>
<td>Expert-Based Instructional Design</td>
<td>Design &amp; Develop</td>
</tr>
<tr>
<td>Research-Based Instructional Design</td>
<td>Deploy &amp; Evaluate</td>
</tr>
</tbody>
</table>

#### Feedback
- Feedback for current project
- Feedback for future efforts

---

**Lydia Ko**

---

**1900 Kite**

**1901 Glider**

**1902 Glider**

**1903 Success!**

**1905 Flyer III**

“Their careful wind tunnel tests produced better aeronautical data than any before, enabling them to design and build wings and propellers more effective than any before.”
# 6

Kirkpatrick Model Is Not Enough!
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)

#7

Smile Sheets Can be Improved!
My Journey in trying
To create a better Smile Sheet

My “New” Smile Sheet
### Specific Concepts

#### How much Value?

**Value of Specific Information**

<table>
<thead>
<tr>
<th>Learning Concept</th>
<th>Circle ONE # Below</th>
<th>Circle ONE Below</th>
<th>Circle ONE Below</th>
<th>Circle ONE Below</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td>How valuable is the concept to you?</td>
<td>How new was concept to you?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value of this Concept (Circle NUMBER)</td>
<td>Low 1 2 3 4 5 6 High</td>
<td>1. Concept was new to me. 2. Deepened earlier understanding. 3. Provided nice reminder. 4. I already use concept regularly. 5. Most people already know this.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>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.</td>
<td>Value of this Concept (Circle NUMBER)</td>
<td>How well was concept taught?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low 1 2 3 4 5 6 High</td>
<td>1. Taught really well. 2. Taught well. 3. Taught inadequately. 4. Taught poorly.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Concept New?

1. Concept was new to me.
2. Deepened earlier understanding.
3. Provided nice reminder.
4. I already use concept regularly.
5. Most people already know this.

#### How Well Taught?

1. Taught really well.
2. Taught well.
3. Taught inadequately.
4. Taught poorly.

### Simple Overall Ratings

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

<table>
<thead>
<tr>
<th>Overall Ratings</th>
<th>Very Little Value 1 2 Average Value 4 5 6 Highest Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low 1 2 3 4 5 6 High</td>
<td>1. Taught really well. 2. Taught well. 3. Taught inadequately. 4. Taught poorly.</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Overall Comfort</th>
<th>Very Uncomfortable 1 2 3 Average 4 5 6 Very Comfortable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low 1 2 3 4 5 6 High</td>
<td>1. Taught really well. 2. Taught well. 3. Taught inadequately. 4. Taught poorly.</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low 1 2 3 4 5 6 High</td>
<td>1. Taught really well. 2. Taught well. 3. Taught inadequately. 4. Taught poorly.</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low 1 2 3 4 5 6 High</td>
<td>1. Taught really well. 2. Taught well. 3. Taught inadequately. 4. Taught poorly.</td>
</tr>
</tbody>
</table>
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.willardworldlearning.com)

Personalized

Your Immediate Impressions of the Learning Experience

Provide your feedback below - I will read each response carefully.

Value of Specific Information

<table>
<thead>
<tr>
<th>Learning Concept</th>
<th>Circle One # Below</th>
<th>Circle ONE Below</th>
<th>Circle ONE Below</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, two common practices in measurement bias our results. Measuring learning at end of program may bias the results.</td>
<td>How valuable is the concept to you?</td>
<td>How new was concept to you?</td>
<td>How well was concept taught?</td>
</tr>
<tr>
<td>Value of this Concept (pace/degree)</td>
<td>1. Concept was new to me.</td>
<td>1. Concept was new to me.</td>
<td>1. Taught really well.</td>
</tr>
<tr>
<td></td>
<td>2. Deepened earlier understanding.</td>
<td>2. Deepened earlier understanding.</td>
<td>2. Taught well.</td>
</tr>
</tbody>
</table>

What else can we do to ensure learners engage our smile sheets?
Performance-Focused Smile Sheets

<table>
<thead>
<tr>
<th>Ultimate Goal</th>
<th>Primary Goals</th>
<th>Secondary Goals</th>
<th>Tertiary Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TRAINING EFFECTIVE? Will the training be effective in supporting on-the-job performance?</td>
<td>Learners Understand?</td>
<td>Learners Engaged</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learners Remember?</td>
<td>Cognitive Supports Effective</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learners Motivated to Apply?</td>
<td>Realistic Retrieval</td>
</tr>
<tr>
<td>MAXIMALLY EFFECTIVE SMILE SHEET</td>
<td>After-Training Supports in Place?</td>
<td>Learners</td>
<td>Spacing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Do Smile Sheet Results Distinguish between Different Levels of Success?</td>
<td>Sit-Action Triggers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inoculated</td>
<td>Job Aids</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Supervisors Follow-up</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Do learners remember enough to 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?**
- **While smile sheets are not capable on their own to determine effectiveness, we should at least try to examine the four goals, (1) understanding, (2) remembering, (3) motivation to apply, (4) after-training support.**
- **We should use smile sheets to send stealth messages to our stakeholders, including senior decision makers, instructors, instructional designers.**
### 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 course topics taught, HOW ABLE ARE YOU to put what you’ve learned into practice on the job?

- NOT AT ALL ABLE to put concepts into practice: Deplorable
- GENERAL AWARENESS of the concepts taught: Unacceptable
- CAN TRY ACTUAL JOB TASKS, BUT NEED MORE EXPERIENCE: Acceptable
- CAN DO ACTUAL JOB TASKS AT A FULLY-COMPETENT LEVEL: Superior
- CAN DO ACTUAL JOB TASKS AT AN EXPERT LEVEL

Percentage of Respondents

A Question About Motivation...

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.
Delayed Smile Sheets

Delayed Smile Sheets offer Great Promise!
Delayed Smile Sheet

Have you used what you learned in the workshop to make a significant improvement in your work?

• No, and I doubt that I will use what I learned.
• No, but I probably will use what I learned.
• No, but I have a plan to use what I learned.
• Yes, I have already used what I learned.

Are You Using What You Learned?

Yes!!

What Enabled You to Apply What You Learned?

No

What is Holding You Back?

Feedback to:
• Management
• Learners
• Us
What is holding you back from teaching others what you have learned?

- The learning content is not relevant to my colleagues’ work.
- I don’t think the learning content is valid.
- I don’t see how teaching others will benefit me.
- I can’t remember the learning content well enough.
- I’m not very good at teaching others.
- I haven’t had the time.
- I have had higher priorities.
- I have not had the resources.
- The risk of teaching this to others is too high.
- I’m still working to persuade management of the value.
- Other, please specify ____________________________

Delayed Smile Sheet

— — — Feedback to Management, Learners, and/ or Us.

Message to Management

“More management-driven than self-driven.”

What are the top 3 to 4 reasons that have enabled you to put the workshop information into practice?

- I remember (or can look up) the workshop’s learning points.
  - I believe strongly in the workshop’s learning points.
  - I have the authority to make the necessary changes.
  - I have decided to lead a change effort to make this happen.
  - I have the time to put this into practice.
  - I have the resources to put this into practice.
  - I will be acknowledged or rewarded if I do this.
  - I will be sanctioned or punished if I do NOT do this.
- My management has made this a high priority.
- My management is fully supportive of my efforts.
- My coworkers are fully supportive of my efforts.
- My Team (or our whole unit) is working together on this.
# 10
Stealth Messaging
Some Stealth Messages We Might Want to Send

1. Training is not enough.
2. Training for awareness is a weak and insufficient approach.
3. Training should specifically minimize forgetting and support remembering.
4. It is better to train a few things really well, than many things perfunctorily.
5. Training should motivate subsequent on-the-job application.
6. Training content must be correct, validated, and relevant.
7. Smile sheets are inadequate.
8. We must measure to provide ourselves with feedback to spur continuous improvement.
9. Prompting mechanisms (like job aids and performance support) should be used—as appropriate—as an integral part of training and as a replacement for training.
10. On-the-job learning should be leveraged in addition to formal training.
11. Learners’ managers and others within the learners’ chain of command should support after-training application.
12. Supervisors play a critical role in enabling the success of training and in helping people be creative on the job.

The Five Failures of Workplace Learning

1. Minimizing Forgetting, Improving Remembering
2. Training Follow-Through
3. Prompting Mechanisms
4. On-the-Job Learning
5. Measurement and Feedback to Spur Improvement

www.tinyurl.com/fivefailures
There are Leverage Points within our SOP's from where we can send Stealth Messages of Recruiting of Work-Learning Professionals, Formal Course Evaluations, Training Requests (our response), Instructor Evaluations, On-the-Job Performance Data, Workplace Performance Assistance, Workplace Performance Analysis, Annual Reports to Management, Course Reviews, and Etcetera.

Does the course use methods that support remembering?
Does the instructor use methods that support after-training follow-through?
Do we determine whether this is a training issue or a workplace issue?
Do we ask trainers to show their consulting skills, not just their platform skills?
Do we utilize Performance-Focused smile sheets & delayed smile sheets?
What messages does a traditional smile sheet send?

What messages does a performance-focused smile sheet send?

What messages does a delayed performance-focused smile sheet send?

Publication Date: 1st Half 2015

To be notified: www.is.gd/WillNews
Performance-Focused Smile Sheets:  
A Radical Rethinking of a Dangerous Art Form

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  
Blog:  WillAtWorkLearning.com  
Twitter:  @WillWorkLearn

Thank you!