

The eLearning Coach Podcast #31
ELC 031: How To Do Course Evaluations The Right Way
With Will Thalheimer, PhD

Show notes: <http://thelearningcoach.com/podcasts/31/>

Welcome to The eLearning Coach podcast, online at thelearningcoach.com. I am Connie Malamed bringing new ideas and tips for success with creating online and mobile learning experiences.

Hello learning people, and welcome to episode 31. In this session I speak with Will Thalheimer, PhD, who is trying to fundamentally change the way that most people in our industry conduct course evaluation. Will is the author of *Performance-Focused Smile Sheets: A Radical Rethinking of a Dangerous Art Form*. He has worked in the learning and performance fields in many different roles for 30 years. Will started his consultancy Work-Learning Research in 1998 to build bridges between research and practical wisdom. Here's the interview.

Connie: Hi, Will. Welcome to the eLearning Coach Podcast.

Will: Connie, it's great to be here.

Connie: I have been going through your book *Performance-Focused Smile Sheets*, and it's really about measuring and evaluating learning, why do you think this is so important?

Will: In brief, if we don't measure we can't improve what we've done. It gives us those feedback loops that help us do our job as professionals, which is to not only put something out there but to see how it works and make it better. So, bottom-line: we don't measure we, don't improve and we're not doing our jobs.

Connie: Good point. What have you seen over the years that inspired you to write the book?

Will: I've been in the training and development field, at least that's what we called it at one time, for thirty years now. I've been a trainer, I've been an instructional designer, so I've gotten my own smile sheet feedback before. And I've also looked around our industry and I've been doing work on research stuff, my research and consulting practice for 17 years, I've worked with a lot of clients. And basically the bottom line is our smile sheets don't work that well for us. It's not just anecdotally what people say, but it's also what you see in the research. So, there's research on over 150 scientific studies done with two different meta analyses, and they found a correlation between smile sheet results and learning results at 0.09, which is virtually no correlation at all. Below 0.30 is a weak correlation, so 0.09 is like correlating how much people eat peanut butter with their eye color, or something like that.

Connie: That is an amazing statistic. In the generic construction design model, at what points do you think it's important for designers to get feedback?

Will: That's a great question, and we have to be a little bit careful because sometimes we think that when we evaluate the ADDIE model, the ease at the end and we should only get feedback at the end, but good instructional designers, or instructional design teams, or eLearning development teams need to get feedback at all stages of the process, not only from learners but also from stakeholders, from learning results, etc. My book is about smile sheets, but one of the things I emphasize in that book is that smile sheets should not be our only methodology. And this can start in the beginning with our needs assessments, I can go through to sort of user testing. Julie Dirksen talks a lot about that, about how important that is that oftentimes in our field we put out some learning, particularly eLearning, we never even watch our users use it.

Connie: Right. We just upload it to the LMS and start on the next project.

Will: Exactly.

Connie: Now let's turn to smile sheets, and just to make sure we're all on the same page, what is your definition of a smile sheet?

Will: I'm glad you asked that because I assumed that everybody knew what a smile sheet was, but I'm finding now as I go to convey the messages of the book that not everybody does know. People call them different things, reaction sheets, reaction forms, student evaluation forms, etc. But in essence it is a set of questions that we give to learners to get their feedback on the training or the learning intervention. I use the word smile sheets even though oftentimes today the questions are delivered electronically, but I use the word because it is sort of the most popular word, at least over on this side of the Atlantic Ocean. I know in Europe they say happy sheets, but it's the most common word so I just use it generically.

Connie: I know you've put a lot of years into researching smile sheets, which implies that there are definitely some problems with the traditional approach. What do you see is the problem with traditional smile sheets?

Will: Well, we know they don't work. We looked at the research and everybody talked to in the industry goes, yeah, well, we've got to use smile sheets, but I don't really put much stock in the answers. But the particular reasons that they're not that effective, one of the biggest reasons is that we are asking our learners for their feedback, and when we do that we really have to support them in their decision making. Because when they're filling out a smile sheet they are making decisions. So oftentimes we like to believe that our learners know learning really well or know their own learning really well. But there's research to show that's not always the case, they make a lot of mistakes in decisions about their own learning, so if we ask them about their opinion of the courses

that we teach, we're going to get some good information and some bad information. Now, this doesn't mean that we shouldn't be asking them, but it does mean that we should be supporting them in that decision making.

So, one of the things that we do that's really bad is we use Likert scales or numeric responses, so on a scale of 1 to 7, 1 to 6, whatever. The reason that like unlike scales are problematic, even though they're the most common thing, we see them in surveys, etc., they're problematic because the answer choices really don't provide enough granularity, so we ask people to choose between 'strongly agree' and 'agree' and 'neither agree nor disagree', etc. And when they're making those choices, there's not enough differentiation between the answer choices to help them make a precise decision. So, one of my recommendations is that we have answer choices that are much more clear.

Connie: Is that what you're talking about when you say giving the support?

Will: Absolutely. And also, there's a lot of biases in some of these questions that are out there, and I think you have some example question to give me.

Connie: Yeah, okay let's do that. In your book you have I think six examples of ones that may be good or may be poor, so let me ask you about this one because I see this one all the time. 'Overall, how satisfied are you with this learning experience?' And then there are five responses that go from 'not at all satisfied' to 'extremely satisfied'. What is wrong with that type of question and response?

Will: There's two things wrong with that. One, they're using a like unlike scale, so the 'strongly agree' to 'disagree' or whatever those kind of terms are. Again, that doesn't give people much of choice. One of the things that happens because of that is that learners go through our smile sheets and circle the same number all the way down the sheet, because they're not really paying that much attention on a specific question. So we've got garbage-in garbage-out problem. We show them questions that they feel are not really getting it what's really important, and they just circle them really quickly so they're not thinking about that. Then we get bad data and we make decisions based on that.

The other thing about that question is asking people how satisfied they are, that satisfaction is probably not a good proxy for the effectiveness of the course. In fact, that's what we should see in those two meta analyses that those kind of questions that we've been asking through the many years are not related to actual effectiveness of the learning.

Connie: Okay, satisfied is a somewhat vague term and it doesn't necessarily have much to do with learning?

Will: Right. And in the book I talk about the four pillars of training effectiveness, and those are: do they understand the material, are they able to remember it, are they motivated to apply it, and are there after-training supports in place. We know that training by itself is not as likely to be implemented as if we had some after-training coaching, reinforcement, etc. So, those four things. So when we're thinking about the effectiveness of our smile sheets, we ought to have questions that relate to that. Now, satisfaction does not necessarily relate to those four things. So, again, that's where we're missing the target in looking at effectiveness.

Connie: What do you think about this one: 'I learned new knowledge and skills from this training'. And then there are seven responses from strongly disagree to 'strongly agree'.

Will: Like unlike scale problem. And that question is leading the witness. Say it again and people will hear that it is a leading question.

Connie: Okay, here we go: 'I learned new knowledge and skills from this training'.

Will: So it's almost begging the learner to say good things about the course.

Connie: One thing I was curious about, you say that smile sheets can't be used in isolation. So what would a holistic approach look like?

Will: One of the things I emphasize in the book is that we shouldn't use smile sheets alone. Now, sometimes we're going to do that, but at least every once in a while, and for our courses or our learning experiences they're really important, we ought to be doing more than that. Now, people are familiar with Kirkpatrick model, so I'll talk in those terms. When we think about the four levels of the Kirkpatrick model there's more that can be done in terms of learning evaluation beyond those four levels.

But, nonetheless, we ought to be measuring people's understanding. Do they understand the concepts, do they have the right mental models about the concepts? I used to talk about Kirkpatrick's level 2.5 as well, and that's decision-making. Can they not only know the material but can they make decisions based on the material. We also want to figure out whether the learners are applying it to their work, whether they are successful in that, whether they have the supports that are needed back on the job, are there managers supporting them, do they have the resources, do they have the time. Because we can't just look at whether they're successful or not, we have to figure out what's causing that success, what obstacles they have, what's enabling them to be successful as well.

Connie: Can you talk about the two goals of a performance-focused smile sheet that you discuss in the book?

Will: There's two that I offer in the book. Number one is does the smile sheet tell us whether the course is likely to be effective as a learning intervention, that's number one.

And number two, do the results of the learning communicate with clarity and urgency to guide our decision-making, to guide our actions as developers. So the first one: Is the course effective? That's where I talk about the four pillars of training effectiveness. And those are: do the learners understand, are they able to remember what they've learned, are they motivated to apply what they've learned, and are there after-training supports in place. Now, all those things are based sort of on the science of learning stuff. In the last ten to fifteen years the research that the academics are doing has really solidified. We now have a very strong base to guide our learning design. But we also ought to use that base to design our learning measurement. And too often we are asking questions that are not related to those things.

Connie: What kind of questions would enable us to get feedback on learning effectiveness? Can you give me an example?

Will: One of my favorite questions is a little bit of a long one but I'll read it off for you if you like. 'In regard to the course topics taught, how able are you to put what you've learned into practice on the job?'

You can see the question stem is already about performance on the job, it's not necessarily about the classroom. I'll just read the first three answer choices. 'I'm not at all able to put the concepts into practice.' That's clearly a bad thing, if you've taken training and you can't put it into practice that would be unacceptable. 'I have general awareness of the concepts taught but I will need more training, practice, guidance or experience to do actual job tasks using the concepts taught.' And the other one, 'I am able to work on actual tasks but I will need more hands-on experience to be fully competent using the concepts taught.'

You'll notice a couple of things. Number one, this question is about performance. That's clearly one of the goals of a smile sheet, do we know whether this course is effective in supporting on-the-job performance? And, number two, you can see there's more granularity between the answer choices. The first one I mentioned, 'I'm not at all able to do it', the second one, 'I have general awareness', and the third one, 'I can do actual job tasks but I need some additional supports, additional experience'. You can see when a learner is thinking about which one to pick, they are sort of able to get their heads around that and say I really do just have general awareness for this. So that's a much better decision-making process than choosing between 'strongly agree' and 'agree'.

Connie: Yes, I can really see that. And I can see how it really gives important feedback to the designer.

Will: Yeah. The second part of my answer was one of the goals of a smile sheet should be to communicate clearly. Normally what happens is we go on a five-point scale, whether we use a five-point Likert scale or a five-point numerical scale, we get a number. So my course is a 4.1, right? Well, a number by itself is pretty meaningless, particularly a number that has come out of some fuzzy adjectives or a fuzzy number

scale. So we get this 4.1, what do we do with it? Well, we could compare it to our previous courses. But, wait a minute, those course were evaluated based on the traditional smile sheet as well, and we know from those meta analyses, that research, that those results are meaningless. So comparing it to our previous results is a meaningless exercise. We could also compare that 4.1 to a standard. I used to run a leadership development product line and we wanted our trainers to get 4.5 on five-point scale, they started to get below 4.0 we would coach them up, and if that didn't work then we would push them out. But that standard was pulled out of thin air as well.

The other thing that's happening these days is there's a bunch of vendors telling us that we need to compare our smile sheet scores to other companies. Well, if they're using bad smile sheets, then we're comparing our numbers to meaningless numbers, so that doesn't work as well. So, having that 4.1 doesn't tell us much. But if you have an answer choice that says, 'I have general awareness of the concepts taught but I will need more training practice, guidance, or experience to do actual job tasks', and people pick that choice, then I've got something and I can share that very clearly with my learning team and with my other stakeholders in the organization. I can say from this course they're having general awareness but they're not really showing that they're able to do actual job tasks. Is that acceptable or is that not acceptable?

Connie: Yeah, so much better.

Will: There's one more thing about the granularity of answer choices that's important. Typically, right now in our courses we get a number like 4.1, my course if a 4.1, and that doesn't tell us very much about what to do. It doesn't communicate to us with any urgency it doesn't communicate to us with any clarity about what we should do. Should we get rid of the course, should we celebrate this course, should we fire the instructor, should we make tweaks, what do we do? Well, 4.1 sort of puts us in a state of paralysis. We don't really know what to do so we don't do anything, so we don't take the time to then improve the course the way we ought to be doing it.

I've seen this with clients. I've got clients who come to me and say, Will, you know, I've hired these subject matter experts to do training for us. And I know that they're not following some of the principles of research based learning design. I know that they're just throwing out a bunch of content and there's not enough practice and there's not spacing and all these things. But, you know, I try to convince them or try educate them and they just won't change. And they come to me and tell me look at our smile sheet ratings, we're doing fine. Now, if you have a question where you have real, clear answer choices, like the one that I suggested, between answer choices talking about a learner having general awareness versus actually being able to do job tasks, that can then create a clear distinction. And then you can show it to the trainers and say, look, you're creating awareness but we really want you to help skill them up. We want them to be able to do something at the end of this program. And that's a much clearer message.

Connie: Yeah. That is an actionable result. In the book you talk about the concept of stealth messaging. What is that and how is that relevant to our field?

Will: It's a great term, isn't it? [chuckles]

Connie: Yes, it really is.

Will: Well, here's the thing, we send messages all the time. When we hire a new trainer, we ask them to come in and show us their platform skills. We ask them to make a presentation. So the message we're sending to them then is what's important is what goes on in the classroom. If we also wanted our trainers to be a little bit of performance consultants, we should also ask them to demonstrate some of their consulting skills. It's all about the messaging that we send. Now, as workplace learning professionals, we have a lot of standard operating procedures, things that we do all the time. One of them smile sheets.

Typically, our smile sheets send messages that what's important is whether the learners are engaged or whether the learners like the material. That's the focus. But we know from the research on learning that that should not be the only focus. We need to dig deeper than that. It would be nice to create a smile sheet that sent messages about what's really important. So the question I shared with you earlier is about on-the-job performance and how people feel about that.

But we can also send messages about motivation. Are you motivated to apply what you've learned? That's a critical thing. If people come to a class, they take our learning and they learned a lot but they're not motivated to do anything about it. They're not going to put it into practice, they're not going to overcome the obstacles, they're not going to take the time, they're not going to make the necessary adjustments you have to make to put something in practice in a way that's successful.

Connie: So what you're saying is the questions that we ask are giving a message as to what we think is important?

Will: Absolutely. And we ought to align those messages with what we know about learning effectiveness, and I'm a big research guy so I think there's a lot of good research that we can align it to.

Connie: I just wanted to clarify two things. One is a lot of the listeners are designing and developing eLearning, and I know your example just then was about standup training, and I just wanted to confirm that you think that these evaluation or smile sheets and the ways of approaching them, it's just as applicable to eLearning too?

Will: Absolutely. In fact I think it's applicable to other types of learning as well. People go into conferences, people in meetings and learning on the job, you could use the same principles in that as well.

Connie: I have one more question I wanted to ask, and that is a lot of the examples that we discussed and that are in the book are written at a high reading level. And I'm wondering how would you approach these kinds of questions if you're writing for something like a sixth grade reading level, and do you think that you can still get helpful feedback if you have to write at a lower reading level?

Will: Yeah, absolutely. The principles that I'm suggesting in the book can be applied at lower reading level. For example, I was just helping a client out the other day and they wanted a question on engagement. You know engagement is a sort of highfalutin word, so instead of writing about engagement I asked a question about how able were you to pay attention. And we had three answer choices. You don't have to have four, you don't have to have five choices, you can lower the number of choices, you can use words that are simple. The ideal way to do this, of course, is to pilot test your questions with the audience, the kind of target audience you're going to use them with. Do think aloud protocols, etc., to find out what they're taking away from those questions. Even just simplifying it and using fewer choices does the trick as well.

Connie: Really good answer. I know I just said we were ending but I really did want to let the listeners hear a little bit about delayed smile sheets. Can you talk a little bit about that?

Will: Sure. One of the things I talk about in my measurement workshops is the fact that people learn, they go up a learning curve, but they also then slide down a forgetting curve. And when we measure at the top of the learning curve we're sort of getting a biased result, as everything is top of mind for the learner and they're just sort of spitting it out. So that's truly related to level two learning, but if you think about it it's also related to level one. So one of the things that we can do is we can also provide learners with a smile sheet, and I typically recommend between two to four weeks after the learning ends to get their feedback then. So that gives us a different perspective, it gives them a different perspective as well when they're answering the questions. It helps them focus on whether they've been able to use this information, it gives us an opportunity to ask whether they've been successful in applying what they've learned to the job. Then we can pivot to a different set of questions, so if they say they are successful we can say, oh, wow, what enabled you to be successful, and we can list a bunch of things. If they say no I haven't yet implemented the training, you can say, well, what are the obstacles that you've been facing. That's really great because we can get feedback to improve our course, number one, if they seem to be facing similar obstacles. Like if they say I can't remember the learning material, then we've got to go back to our course and design it to help support remembering. They tell us we've been successful because my manager is really supporting me in this, we can say great.

But we can also make perhaps mid-course corrections. I've done this with my workshops. I did a delay smile sheet, found out the reason some people were being successful and said, hey, here's the reason your people are being successful. I

remember one where the reason people were being successful was because their manager was really pushing this. So I said don't stop pushing it, because if you stop pushing it they may stop doing. The same thing with obstacles. If you find out people don't have the time or the resources then you can go back to management and say this is what we're hearing, if you want people to apply this new skill that we've taught them, we remind you to do [some of the? 23:18] different things.

Connie: That's really good to know. Will, thank you so much for your time.

Will: Connie, it's been a pleasure. Thank you very much.

I really enjoyed that conversation and I hope you did too. For the show notes where you can find a link to Will's book, which I highly recommend, go to thelearningcoach.com/podcasts/31. And, if you would, please leave a review for this podcast on iTunes. That's it for now. I'll talk to you next time. Take care.