EQUITABLE EVALUATION

VALIDITY EXPLANATION

WHAT WE DID AND LEARNED
December 2018
As part of the Equitable Evaluation Design Sessions (hosted by the Colorado Evaluation Network, sponsored by the Colorado Health Foundation and the Colorado Trust, and facilitated by Jara Dean-Coffey), Vantage Evaluation took on two design challenges:

1. How can we build personal awareness of inequities and examine the professional implications of these inequities in our evaluation work?
2. How can we expand the definition of validity in evaluation used by nonprofits and foundations in our community?

In this summary, we will focus on the second design challenge: why Vantage Evaluation decided to focus on expanding the definition of validity, what we did (or tried to do), what we learned, and where we are going next. Our hope in sharing our experiences is that others might find inspiration to engage in similar work, and that those organizations will take what we did and improve upon it. The first design challenge, related to personal and professional awareness of equity, is the topic of a separate summary.

**WHY THIS IS IMPORTANT FOR THE EVALUATION FIELD**

Our decision to focus on expanding our target clients’ definition of validity emerged from a realization of the connection between historical definitions of validity and structural inequities. Traditional notions of “valid methodologies” reflect white, male, cis, middle-class ways of knowing. If the evaluation field does not broaden our definitions of validity, we will continue to perpetuate inequities in the ways we design “rigorous” work. There are two pieces of this puzzle: the demand side and the supply side. Conversations about redefining equity have already started among evaluators (the supply side), most notably in the work of Karen Kirkhart on multicultural validity. These conversations are primarily theoretical and have not yet gained broad prominence in the field. Because evaluation is a service business, we believe that no progress will be made until the demand side (clients) re-visits what counts as valid. So that’s where we focused: how can we take a theoretical conversation about defining validity among evaluators and bring it into practice with our clients and potential clients?

**WHY VANTAGE?**

Vantage Evaluation takes pride in our ability to translate evaluation concepts into language that resonates with our clients and potential clients. We seek to be the bridge between professional, academic, and theoretical evaluation, as well as the ways those concepts play out in the day-to-day reality of the nonprofits and foundations with which we work. Additionally, we get feedback from clients that our ability to do this well is part of what sets us apart. We thought we might be able to use those skills in service of broadening the definition of validity.

**WHAT WE SET OUT TO DO**

Our plan was to use our position of privilege and power to develop strategies and language for reframing validity in a way that is meaningful for our nonprofit clients and potential clients. The goals of this work were to:

1. Strengthen Vantage’s ability to articulate a broader understanding of validity.
2. Use our power and position to shift the conversation around validity away from purely methodological considerations in Colorado, indirectly raising awareness of inequities in the field’s current evaluation priorities.
3. Test language reframing validity and identify messages that resonate most with nonprofits, foundations, and other evaluations (contribute to the field).

We set out to test language during three different times (potential client meetings, launch meetings, and design meetings). After each use of the language, we planned to debrief about how the client reacted, what questions they had, if the language resonated with them, etc. After that debrief, we planned to revise the language and test it again. The goal was to have tested and standardized language for talking about validity at the end of the project.

We scripted language for four time points to talk about validity:

| Introductory Conversation With Potential Clients | [After an introduction to why evaluation is different than research] One of the ways evaluation has continued to evolve away from traditional research practices is in the way we think about what makes a finding valid. In research, validity was a product only of research design. If the design was strong, the results were trustworthy. In evaluation, we’ve recognized that strong research design is not the only way to demonstrate that something is true. We see this in our everyday lives—for example, the research on climate change is solid, but people still do not agree that climate change is true. Often, it is because someone they trust has said that climate change is not true: this is relational validity, where a finding is “true” because of the relationship with its source. We also see this in nonprofit programs. We had a client recently for whom we built a solid research design, but when the results came back, they did not align with the staff’s experience of the program—in other words, “experiential validity.” As evaluators, we work hard to keep an eye on all of the different ways that people know, and thus are able to generate findings that are valid, believable, and actionable to a more diverse group of stakeholders. |
| Evaluation Orientation Trainings | Just like we all have different learning styles and thinking preference, we all have different ways to conclude that something is true. There are five core ways of knowing: First, some people are only willing to say, “Yes, I believe that fact,” with hard scientific evidence that is methodologically sound. Second, some people believe something when it aligns with their lived experience of the world. Third, others believe a fact when it comes from a trusted source, even if it has neither scientific evidence nor alignment with their experience. Fourth, still others believe that facts are true when they are presented in a way that holds a mirror up to challenge our prior understandings within the specific context. However, we have found that by only focusing on this one way of knowing, the findings from the evaluation do not always resonate with our clients and, as a result, the evaluation is not as useful to them. We think this is because of the history of evaluation: information was presented as “valid” if (and only if) it came from a rigorous research design. Because of this, clients are tricked into thinking that is the only way to know, and don’t realize it’s not until the end of the project when they are presented with results that don’t resonate with them. So, we have this conversation at the beginning of the project. As you are thinking about what you want out of this evaluation effort, also think about how you will know the findings are true at the end of the project. What do you need to be confident in the findings? We will pick this conversation up again at the launch meeting. |
| Project Launches | Just like we all have different learning styles and thinking preferences, we all have different ways to conclude that something is true. There are five core ways of knowing: First, some people are only willing to say, “Yes, I believe that fact,” with hard scientific evidence that is methodologically sound. Second, some people believe |
something when it aligns with their lived experience of the world. Third, others believe a fact when it comes from a trusted source, even if it has neither scientific evidence nor alignment with their experience. Fourth, still others believe a fact when the theory underlying the evidence matches the cultural setting. And fifth, some believe that facts are true when they are presented in a way that challenges our prior understandings within the context and consequences of the fact. In the same way that we don’t force a certain set of questions and methodologies on our clients, we also don’t force a certain way of knowing and believing onto them. However, we often find out which ways our clients believe something to be true after the end of our evaluation. We think this is because of the history of evaluation: that information was presented as “valid” if (and only if) it came from a rigorous research design, so clients are tricked into thinking that is the only way to know. Then, they get those rigorous research results, and they don’t match their experience, so they don’t believe them. We are trying to do a better job of having these conversations on the outset. So, in thinking about the different ways of knowing, which resonates more with you? Is it a combination? If so, what percentage of the pie does each approach reflect? We’ll be sure to tailor our design to make sure we get evidence in the end that you can conclude is true.

Just like we all have different learning styles and thinking preferences, we all have different ways to conclude that something is true. In this evaluation design, we’ve incorporated elements to support many different ways of knowing. Our goal is to generate findings that are valid, believable, and actionable to as diverse a stakeholder group as possible. For example… [pick and choose the ones you’ve used in the evaluation design]:

- **Methodological Validity:** For stakeholders who believe something to be true because it comes from sound research design, we’ve followed best practices by including ____ and staffing the project with expertly trained Masters and PhD level evaluators.

- **Experiential Validity:** For stakeholders who believe something to be true when it matches their experience, we have added a phrase where we share back the results with the program participants to make sure we have captured the experience accurately.

- **Relational Validity:** For stakeholders who believe something to be true when it comes from a credible source, we have worked incredibly hard to build a reputation as a trustworthy evaluation partner in Colorado and will continue to build trust via ____ and ____ throughout this project.

When people think about generating valid findings, often their brains go straight to things like sample size. In other words, methodological considerations that academia had taught them to associate with good science. However, there is more to valid findings than methodological designs: things like which voices are represented, how accurately they are represented, if the data collector is trusted, if the report is trusted, etc. The design we present to you today incorporates multiple ways of ensuring valid findings, such as: [highlight additions that add validity, including methodologically, but also things like member checking, advisory committees, cognitive interviews, etc.]

**WHAT WE LEARNED**

We were able to test validity language at two different time points: one with a potential client, and one during an Evaluation Orientation Training webinar at the start of a project.
Introductory Conversation with Potential Client: Fail. In a first conversation with a potential client, we intended to introduce the concept of a broader conceptualization of validity. But in the moment, we got so excited about talking about the project and program and organization and the team and how evaluation would be so great to help them achieve their goals… that we completely forgot about bringing up validity. Oops.

Evaluation Orientation Training: Who Knows? Before the start of a new project, we hosted an evaluation orientation training webinar for the client and evaluation advisory committee. We incorporated a slide into our existing content and the Evaluation Orientation Training script. This content was added at the beginning of the section discussing matching methods to key evaluation questions. Evaluators on the project reflected that this multicultural validity content flowed nicely with our existing content, and could potentially encourage clients to think about different ways of knowing as they are identifying potential evaluation methods. However, because this training was a webinar, rather than in person, we could not gauge the reaction of participants to this content. We could not see participants to understand nonverbal reactions, nor were there any specific questions or comments about the content.

We had planned to test the Project Launch language with the same client, but were not able to fit that conversation into the agenda for the meeting. During that meeting, these conversations about multicultural validity promoted us to try a Power Mapping activity with meeting participants. During this activity, we had meeting participants identify all of the stakeholders involved in the program, then indicated how much impact the program had on them and how much control they had over the program. This process allowed us to identify stakeholders with high control and high impact to prioritize in the evaluation, as well as stakeholders with a lot of impact but little control or a lot of control but little impact. This helped evaluators and the client understand the different voices involved in the program and the power dynamics that exist among stakeholders.

We learned plenty of lessons through undertaking this validity explanation project this summer:

1. **Scripted does not work.** We started the project with a set of scripts for how to talk about multicultural validity. However, we quickly learned that validity is not a “plug-and-play” type of conversation where we can easily drop something that is pre-scripted into the flow of the conversation. Every conversation is different based on the background, context, and experience of the person or people we are speaking with.

2. **Very hard to change our spiel.** Our conversations and meetings follow a generally practiced flow. As a result, it is hard to break that flow for new material.

3. **Need to understand the theory well before you can translate.** As we got further into this process, we realized that we did not understand all of the components of multicultural validity well enough ourselves to be able to translate it effectively for our clients. For example, we did not (and still don’t) have a good understanding of theoretical validity, and could not provide examples of how it played out in evaluation projects.

While these were important lessons for us to learn, we have also reflected that life happens. Our team had other personal and professional priorities that took up the time and mental energy we were hoping to dedicate toward this project to do it well.
WHAT COMES NEXT

We aren’t exactly sure. Our main takeaway from this design challenge is that it’s really hard to reconceptualize an established concept and then explain it to people with a pretty superficial understanding to begin with. But we’re not giving up, and we hope others will join us and share what they learn along the way, so we can make steady progress together.

At the second design session, we held rotating mastermind groups to troubleshoot our biggest Equitable Evaluation challenges. Elena got some great ideas from other members of our cohort. Specifically, we think we may have been trying to introduce validity too early. It was too hypothetical for our clients and prospective clients. With all evaluation work, it might work better once specifics are in place. So, here are the next things we are going to try:

- Go through our usual process to help the client establish key evaluation questions
- Ask them what they think the answer to each key evaluation question
- Then ask, if the answer was the opposite, what would need to be in place for you to believe that answer?
- Then, use what they say as a jumping-off point to discuss multiple valid ways of knowing!

We’ve been starting to incorporate this new conversation entry point into our work, and it seems to be working much better.

NOW IT’S YOUR TURN

We tried this at Vantage— now it’s your turn! Take what we did and our lessons learned, then start to have conversations about multicultural validity and different ways of knowing. We’re happy to help you think through how you might get started!

Next, let us know how it went! What did you try? How did your audience react? What would you try differently next time?

Don’t forget to connect with the Equitable Evaluation community or on social media using #EquitableEval.

Good Luck!!

Elena Harman

elahman@vantage-eval.com

Laura Sundstrom

lsundstrom@vantage-eval.com