

Research Questions	Validity Threats	How is Validity Threatened?	Strategies to Address Validity Threats
How do AP Statistics teachers understand statistical inference?	Short acquaintance with the teacher participants.	Teachers may not feel comfortable to be truly forthcoming about their knowledge or lack of knowledge with a relative stranger	Participant selection can aid in ensuring reasonably competent and experienced teachers are chosen. I intend to go through the College Board to obtain access to AP Statistics teachers who serve as readers for the AP Statistics exam. These teachers must have taught AP Statistics for at least 3 years.
	Teachers may be defensive if they interpret my inquiry as judgmental.	Teachers may not want to admit to any information which they feel casts them in an unfavorable light	Teachers will be approached as professionals who have the expertise and knowledge which my study needs. Total anonymity will be guaranteed. In interview sessions, only the teacher and I will be present; so, there should not be any feeling of peer pressure. Questions will use non-judgmental wording, such as, “tell me about a time when ...” or “what is your opinion of the most important foundational ideas ...” Teachers are not be judged or scrutinized; they are being asked for their expertise and their insight.
	Teachers may incorrectly assess their own understanding.	Teachers may know more or less than they think they do.	I will ask teachers to give me an example of how they teach a certain lesson. What do they include and do not include, and why?
	Researcher Bias - My understanding and teaching of statistical inference	Interjecting my own opinions and my way of doing something is not the focus of my study, and, doing so would probably stifle the teacher responses. Some may also be insulted.	I need to keep my mouth shut except when I am asking the questions. No opinions, no agreement or disagreement, no alarmed facial expressions, no interruptions from me. I wouldn't be <i>absolute</i> about this; at a minimum, you want to convey what someone (Weiss?) called "empathic neutrality." However, it's a good general principle. I do not want the teachers thinking that I know more than they do or thinking that I think I know more than they do.
	Reactivity - Teachers may report what they think I want to hear, or what they think paints them as superlative teachers, instead of what they really do in the	Information is not true	There is only one researcher in this study; so, controlling for differences between researchers is not a concern. Eliminating the actual influence of the researcher is not possible (Hammersley & Atkinson, 1995). I need to present myself as a person who is requesting information from subject matter experts who are both experienced classroom teachers and AP exam readers. I am an AP Statistics teacher; but, I am not an AP Statistics exam reader; they know more than I do. Even more importantly, they know what I want to know.

	classroom.		
<p>How do teachers incorporate the concepts of statistical inference into their AP Statistics curriculum?</p> <p>What level of importance do AP Statistics teachers place on statistical inference?</p>	<p>Teachers may be influenced or intimidated by peers in focus groups or by their perceptions of my expectations.</p>	<p>Teacher comfort level, teacher willingness to be forthcoming, and teachers' perceived expectations of peers – limits the information shared</p>	<p>Establish friendly, non-threatening atmosphere</p> <p>Use major themes from interviews as these are likely important points with all AP Statistics teachers.</p> <p>Draw teachers in to discussions of particularly successful lessons, particularly memorable lessons, or personal strategies.</p> <p>Ensure that everyone talks and is recognized. Do not let anyone dominate the conversations.</p>
<p>Is there any correlation between the teachers' understanding of statistical inference and their students' performance on statistical inference questions on the national AP Statistics exam?</p>	<p>Individual student scores are privacy protected; however, a teacher is free to discuss his/her class mean score.</p> <p>Teachers may be willing to state how their students perform overall; but, this will either be accurate or inflated. If a teacher's mean score is low, he/she may feel demeaned or threatened.</p> <p>Some students may have received outside tutoring (which is currently a common practice); this could affect their exam score</p>	<p>I may not be able to get the data to answer this question. I consider this question to be secondary to the others and I am not willing to jeopardize the data collection for the first questions. If data for this question cannot be obtained without posing any discomfort, trust issues, or cooperation from the teachers, then it will be abandoned and left for another day.</p>	<p>I cannot assess if any correlation exists unless I have some exam score data. Some teachers may bristle at being asked for scores. Some may simply not want to discuss it. I think that a reasonable approach would be to ask the teachers what questions on the AP exam their students reported as being most challenging and least challenging. The latter gives them a chance to brag about their teaching and/or their students' understanding. Also, I should ask what concepts the teachers themselves found the students found most challenging in the classroom.</p> <p>To avoid the brutality of asking how their students performed on the exam, I will approach the idea of how their students performed against the national average. I think that each teacher would have at least one question for which their students outperformed the mean. I will ask them to discuss why they think their students did so much better. Then, I can gently and non-judgmentally segue into why they think their students did not do so well on another conceptual area. There could be (and probably are) a host of reasons which cast no shroud of inferiority on the teachers or their students, e.g., a flu epidemic caused everyone to miss two weeks of school; a bad storm caused the power to be out for over a week; the teacher was hospitalized; the textbook was confusing to the students; inaccessibility of study materials; students do not have Internet access at home; or, lower level students demanding too much repetition and causing the teacher to move too slowly to cover all the material. But, hopefully, some there will be some pedagogical reasons addressed also.</p> <p>There will always be some teachers who love to brag about how well their students perform on the AP exam. I will capitalize on any of these opportunities by asking the teacher to tell me more. Also, because the participants are experienced AP Statistics teachers who are at a competency level to grade the national exams, I will be very surprised if</p>

			there is not a significant amount of pride in their work and the accomplishments of their students. But, that is my perception because of the demeanor of the AP Statistics colleagues I know.
--	--	--	--

All truths are easy to understand once they are discovered; the point is to discover them. Galileo Galilei

I decided to do a validity matrix and memo for this assignment for two reasons. First, I find it challenging to try to list threats to the validity. It is true that some stick out like a sore thumb. For example, in my study, anticipating that teachers may become defensive if they find my questions to be judgmental is not an especially shrewd insight. It just makes sense. However, other validity threats are far from this obvious. They need to be ferreted out. This takes a clear-headed analysis of participants, a realistic self-evaluation, and, a comprehensive scrutiny of the design elements. I need to be ever vigilant that my own biases do not cloud my study design or implementation. However, I do not want to send my own knowledge and experiences on vacation for the duration of this study. ☺ My knowledge of AP Statistics enables me to ask questions, and follow-up questions, which a person who does not teach AP Statistics could not. My sketchy part here is that I cannot assume that I understand what a teacher is saying simply because s/he is discussing statistics. I know what a teacher means when s/he refers to a sampling distribution or z-scores. But, the meaning of “students make more precise graphs with means than they do with proportions,” or “students understand Normal curves better than Chi-Square” is not clear. I get the general idea; but, I would need to ask the teacher to elucidate so that I get a clear picture of her/his true meaning. ☺ In my perception, both of those statements are true for my own classes; but, another teacher may say the same thing, but mean something very different than I do. Secondly, I place a tremendous value on validity. The most elaborate study is really useless if there is no reasonable validity established for the conclusions. It makes no logical sense to spin my wheels with study design, data collection, and analysis, if it is all for naught. I do not mean that I do not reach the conclusions I had hoped to achieve; I mean that my conclusions lack validity. Reaching unexpected conclusions is a useful result and I would learn something and be able to contribute something from that. Not what I expected, but valid? That’s ok with me. I found something. ☺

I started with my design matrix from the second assignment and have attempted to expand, refine and develop my thoughts, focusing on validity. This time I included my own possible biases, participant selection, and Maxwell’s (2013, pp. 125-129) checklist. As I have been ruminating with my ideas of potential threats to validity

over the past weeks, I have come to see how pervasive this concept needs to be in the design process. ☺ It is not merely a box which needs to be checked off, such as a pre-vacation list of stopping the mail, getting a dog sitter, etc. Ok, I thought about validity, check, all done. Screeching halt!

Irwin (2008) suggests that different types of data provide different viewpoints and are not confined to distinct research questions; they are different dimensions of the questions. In the same vein, identifying validity threats are not one-dimensional. We want validity attached to every aspect of our research; validity needs to be a dimension of all of our research questions. I would say, a property of the *answers* to our research questions. One of my first thoughts on validity for my study was the perceptions of the teachers. ☺ I wanted them to feel comfortable and know that I was grateful for their honesty and appreciated their expertise. I do not want to appear as a blathering sycophant but I do not want to appear too authoritarian, either. It made me think about just how I could accomplish that (the grateful, sincere, and earnest researcher) while still setting the stage for meaty, edifying, albeit mostly one-sided, conversations. I thought about my own experiences as an interviewee ☺☺ and remembered being both almost embarrassingly loquacious as well as being guarded and reserved in my responses. There were numerous variables which contributed to these scenarios, e.g., the interviewer's attitude and appearance, the comfort of the setting, etc. But, the most memorable characteristic was what I believed the interviewer was trying to accomplish. ☺ Is she trying to get my friend's job abolished? Is someone being investigated? Are they trying to justify more funds for our department? Needless to say, my perceptions drove the amount of information which I was willing to share.

With that in mind, the idea of ensuring that my teacher participants know that the goal is to improve AP Statistics teaching, and that that is the only goal, grew in importance. ☺ So, then, my approaches to ensure this result also need to be expanded. I think that I have addressed my approach, wording of questions, and setting the atmosphere. However, how could I put some sort of control on the teachers? I'm not sure what you mean here. Choosing only teachers who are readers for the AP exam would weed out teachers who are inexperienced or not terribly interested in AP Statistics. It would also, theoretically, ensure some level of passion for the subject. The College Board restricts high school teachers to serving as readers to at most six years. After six years, they are either retired or promoted to table leader (in charge of 14 readers). I do not want to access any Table Leaders because these people are not representative of AP Statistics teachers. But you've already done this by not including inexperienced or unmotivated teachers. (However, further study of Table Leaders may lead to interesting

comparisons.) Selection of AP exam readers ensures a high probability of teachers who seek out PD to improve their practice and are genuinely interested in improving AP Statistics teaching. I also took Dr. Maxwell's suggestion and removed the knowledge assessment questionnaire for the teachers, as this would probably only exacerbate any teacher defensiveness. ☺

I had become confused about the order of focus groups and individual interviews. On one hand, having the focus groups first would invite the teachers to get to know each other, discuss the topic and become comfortable with the study. Holding focus groups first and then individual interviews will aid in creating a safe and comfortable atmosphere of professional inquiry for the teachers. Listening to other teachers' ideas may get them thinking more reflectively about their own practices. It may trigger some memories. This may pave the way for more productive interviews. Then again, if individual interviews are held first, then the ideas which the teachers relate would truly be their own. An important issue. The salient ideas from the interviews would form the core of the focus group questions. The teachers would already be familiar with the study and would have already discussed their own philosophies with me. So, a clear validity threat to the accuracy of information is the potential influence teachers may, inadvertently or not, have on each other if the focus groups were held first. ☺ One teacher's opinion may sway another's. This validity threat will be avoided by conducting the interviews first. Of course, any method has its own characteristic biases. Using only one method to assess a phenomenon inevitably leads to limited and biased results (Greene, 2008). Conducting both focus groups and interviews, coupled with my observations during the interviews and focus groups, will provide varying viewpoints.

There are more threats, I just know they are hovering out there; but, I do not know what they are yet. I can see that some may make an appearance at any time, during my question writing, during a focus group, during an interview, anytime at all. But, at least by starting to memorialize my thoughts on the validity threats which I do recognize, I have begun to think more about these threats and to be mindful of, actually, to be on the active lookout for, others. ☺ Simultaneously, I am thinking about what I would need to do or to design differently or additionally in order to address these threats. ☺ It is like standing on the shore and throwing the starfish back in the ocean. Yes! I will do the best I can, knowing that I will never get them all. But, I'll address every validity threat I can reach.

Mimi:

Excellent work on this; I really don't have many substantive comments.

One thought: this study is a good example of the lack of isomorphism between research questions and interview questions. You want to know how the teachers understand statistical significance, but asking about this directly will probably make them at least somewhat defensive. You can learn this just as well (maybe better) by asking them how they *teach* statistical inference.

References

- Greene, J. C. (2007) *Mixed methods in social inquiry*. San Francisco: Jossey-Bass.
- Hammersley, M. & Atkinson, P. (1995). *Ethnography: Principles in practice* (2nd ed.). London, England: Routledge.
- Irwin, S. (2008). *Data Analysis and Interpretation: Emergent issues in linking qualitative and quantitative evidence. Handbook on Emergent Methods*. New York: Guilford Press.
- Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Los Angeles, CA: Sage Publications.