AAAS Self-efficacy for Public Engagement with Science (PES)

Scale for Scientists

What is self-efficacy? The current scale measures scientists' self-efficacy for public engagement with science, which we define as beliefs in their ability to prepare for and conduct PES activities that created reciprocal interactions with publics. Bandura (1997) defined self-efficacy as, 'people's beliefs in their capabilities to produce desired effects by their actions' (p. vii). A tenet of social cognitive theory is that self-reflection allows individuals to assess their knowledge, experiences, and thoughts as a means of determining their likelihood of success. Bandura maintained that, through self-reflection, people appropriately alter their thoughts, behaviors, and motivation to engage in or continue with various undertakings. The construct of self-efficacy has been studied in many contexts to confirm the relation between ideas about self, behavior, and success. Within the context of science communication, self-efficacy is also a primary motivator among scientists who participate in PES activities (Dudo, 2012; Dudo & Besley, 2016).

How can you use this scale? The Self-Efficacy for PES Scale was designed with multiple purposes in mind. The scale can provide baseline data to describe the range of self-efficacy found among a group of scientists. If embedded within the context of a scientist training program, the scale has the potential to serve as a reflection tool for scientists by providing data on the relative strengths and areas of improvement in their PES activities. The scale can also be used as a longitudinal measure to document changes in self-efficacy over time. Using the scale in this way is advisable only in situations that involve a sustained intervention with multiple interaction points that take place over weeks or months.

How were the items developed? Twenty-five scientists participated in think aloud interviews to provide response process evidence to support the use of specific survey items (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). Scientists were from a range of disciplines. All had been involved in PES within the past year. Most had recently engaged in some sort of public dialogue event (n=16), while others were involved with a university or cooperative extension activity (n=5), policy deliberation (n=2), or knowledge co-production activity (n=2). The think-aloud interviews narrowed down the number of items from 30 to 19 that were intuitive to scientists and yielded a range of rating responses. These items were then administered to a test sample of 297 scientists who had conducted at least one PES activity in the past year.

What is the internal structure of the scale? For the purposes of the current study, graded response models were conducted (see Samejima, 1969; 1996) using MPlus 7 (Muthen & Muthen, 1998-2012). Two models were conducted with all 19 items initially. The first model generated estimates for both the slope and location of each item in the model. The second estimated the location only, and held the slope constant (and thus was akin to a Rasch model). The two models were then compared using a likelihood test; results indicated that the default graded response model was the better fit for the data compared to the model that constrained the slope of each item and thus confirmed that items in the model were differently discriminating with regard to self-efficacy (df=18, p<.001).

For the remaining analysis, the difficulty and discrimination estimates were calculated from the graded response model and then used to determine the items that provided the best measures of self-efficacy. A high discrimination parameter value means that the probability of a correct response increases more rapidly

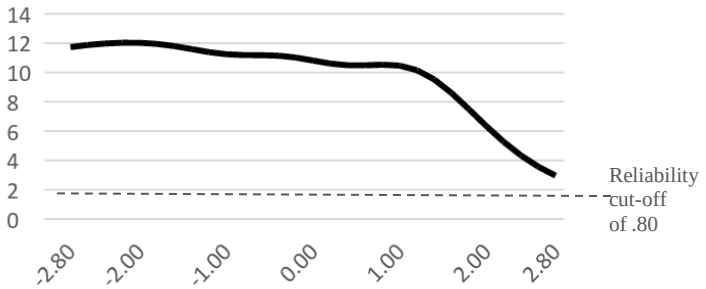
as the ability (latent trait) increases (An and Yung, 2014). Acceptable discrimination values are greater than 1.0; the discrimination values for items on the Self-Efficacy for PES scale ranged from .87 to 2.04. Two items were eliminated due to low discrimination levels.

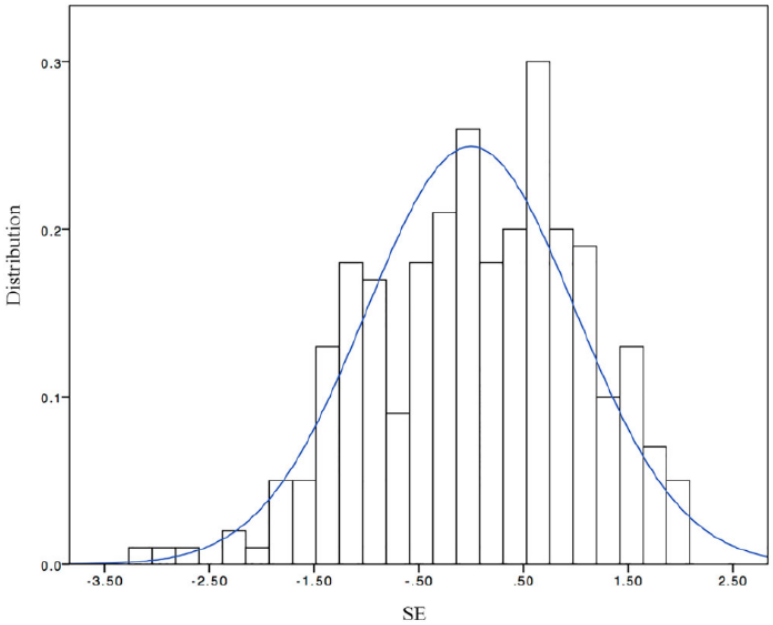
The Item Characteristic Curves (ICCs) were then explored for the remaining items and a pair of related criteria were used to determine suitable discrimination for the purposes of the current study. Items were retained if at least five of the six rating responses were used by scientists to rate their self-efficacy *and* if at least five of the six had the highest curve for some portion of the distribution across self-efficacy ratings. Four items were eliminated because they did not meet these criteria. The final scale includes 13 items; means, standard deviations, and estimated parameters for each are presented below.

			Thresholds					
	Mean	**SD**	1	2	3	4	5	**Discrimination**
I am able to create props/activities/demonstrations that participants find engaging.	4.77	1.14	-3.97	-2.98	-2.16	-0.80	0.95	1.17
I have a hard time finding PES topics that people connect with.*	2.15	1.30	-4.14	-2.52	-1.55	-0.83	0.24	1.28
I find it difficult to leave time for people to share their perspectives during PES activities.*	2.74	1.25	-5.37	-2.57	-1.15	0.01	1.86	1.01
I have a hard time finding the right words to convey my message during PES activities.*	2.47	1.22	-4.76	-2.78	-1.29	-0.50	1.25	1.16
I am good at thinking together with PES attendees about science topics.	4.74	.95	-3.99	-2.79	-1.89	-0.54	1.15	1.78
I am good at knowing when to inform and when to listen during my PES activities.	4.57	.99	-3.47	-2.61	-1.76	-0.32	1.52	1.65
I am able to figure out how to improve PES activities based on the kinds of questions the public asks.	4.71	1.01	-3.29	-2.50	-1.81	-0.49	1.11	1.83
I am able to engage in critical discussion about science topics with non-scientists.	4.89	1.00	-3.35	-2.49	-1.88	-0.78	0.78	1.95
I am able to moderate discussions with participants, even when they include a wide range of perspectives.	4.48	1.03	-3.20	-2.50	-1.41	0.01	1.30	1.91
I am good at reading the audience during PES activities, and making adjustments as needed.	4.60	1.09	-3.00	-2.06	-1.48	-0.33	1.07	2.14
I am good at finding ways to approach difficult topics.	4.62	.99	-4.14	-2.81	-1.71	-0.26	1.27	1.70
I have a hard time answering questions from non-scientists in ways they understand.*	2.28	1.13	-3.67	-2.57	-1.51	-0.58	0.84	1.68
I am able to moderate discussions that allow participants to engage with me and with each other.	4.51	1.01	-3.73	-2.26	-1.57	-0.11	1.36	1.99

*These items are negatively worded, therefore lower scores indicate higher self-efficacy.

The results from the initial graded response model reduced the scale to 13 items that had high reliability from the lowest end of the range up through approximately two standard deviations above the mean (see figure at right). The means and standard deviations for each individual item are shown in Table 1, along with their threshold and discrimination parameters.





The mean Self-Efficacy for PES score for the sample was 4.64 (*SD*=.69), with a range of 2.46 to 6. These results indicate that scientists had moderately positive self-efficacy for PES overall. The distribution of scores indicates that the scale detected a broad range of self-efficacy among scientists. The distribution of scores overall is presented in the figure to the right. Compared to the ideal normal distribution curve on the graph, scores on the Self-Efficacy for PES over-represented on the lower side of the graph and slightly under-represented for higher scores on the scale.

For more details on this analysis, see:

Robertson Evia, J., Peterman, K., Cloyd, E., & Besley, J. (2017). Validating a scale that measures scientists' self-efficacy for public engagement with science. International Journal of Science Education - Part B: Communication and Public Engagement. Retrieved from http://www.tandfonline.com/doi/full/10.1080/21548455.2017.1377852.

For more information contact Karen Peterman, karenpetermanphd@gmail.com.