

Responses to OMB questions about NASA's Evaluation of Summer of Innovation

1. On page 10 of the supporting statement (Part A), NASA discusses hanging negatively worded questions to positive (or deleting). Please provide more information about your rationale – we are puzzled. NASA cites literature describing whether to use a neutral middle option of not – which is an important issue. However, it is unrelated to whether the question is worded positively or negatively. So, there are two issues here: whether or not a list of options should be positive and negative mixed and whether or not a scale should have a middle option. They both affect how one writes a question but they really are two different issues.
 - a. **Mixing positive and negative options together should be done for two reasons. First, it is a way to keep people from straight lining a list, which we believe would be especially tempting for kids with low levels of literacy, for example. Second, it's also done for statistical reasons. In order to properly create a latent variable, one is supposed to measure the construct from both sides and then flip one set in analysis so they're going in the same direction. So mixing the options is both a social and a statistically preferred method of asking a long list.**

We had adapted the pilot survey items so that they were all positively worded as we received reports that students struggled with the negatively worded items. However, recognizing that the literature continues to debate this issue, we have revised the survey so that both positively and negatively worded questions are used.

The Career Interest in Science and Leisure Interest in Science scales were taken from the *Test of Science Related Attitudes (TOSRA)* which includes both positively and negatively worded questions; we have revised the questions back to their original TOSRA wording so that both positively and negatively worded items are used. The *Attitude Toward Science* scale was taken from the *School and Social Experiences Questionnaire*; this survey only uses positively worded questions. Given your recommendation, we have modified a few of the questions drawn from this survey to be negatively worded rather than positively stated. This also ensures consistency with the revisions made to the other two scales. We have also added in any items that we had dropped from those original scales and included them in their original state and will reexamine their results after this year's survey administration.

- b. Whether or not to include a middle option depends on what construct one is measuring and whether or not NASA wants to force people into an answer. When people have little or no opinion about something, they will often pick the middle option. We usually recommend no middle option but adding a "don't know" option if relevant.**

We appreciate your recommendation to eliminate the middle option on our scales and recognize that that the literature has yet to reach a consensus on the issue and can see how in many cases it may be preferable to force people to answer. Following your guidance, we have revised the teacher "comfort" items originally drafted with a five-point scale to now use a four-point scale and have thus eliminated the middle option.

For the purposes of the student survey, however, we would like to retain the middle item. The majority of our items are taken from the *Test of Science Related Attitudes (TOSRA)* survey, which was developed using a five-point scale. So that we may rely on the psychometric properties derived from the testing conducted by its developers, we would

like to keep the scale intact. Furthermore, a “neutral” response is meaningful within the context of this study as we are focused on assessing whether any change occurs in student interest in science: a move from a neutral to a slightly positive or negatively response could provide early indicators that could help identify promising practices appropriate for more rigorous examination in the future.

Several items included in the student survey are taken from The *Attitude Toward Science* scale of the *School and Social Experiences Questionnaire*, which uses a four point scale. However, we want to use a five-point scale for these items so that they are consistent with the scale construction of the TOSRA questions. This consistency is important for two reasons: (1) it may prevent students from confusing the scales on the survey and (2) it allows us to “mix” items, so that we can intersperse the items measuring different constructs (for example, a career interest question could be used before and after a science attitude item) to prevent students from straight-lining the survey.

- 2. On page 11, Supporting Statement A, question 10, NASA needs to say that the information will be protected “to the extent provided by law.”**

We have added this clause to this section.

- 3. Page 3, Supporting Statement Part B introduction, does the first sentence of the second paragraph mean that teachers at the awardees sites are the only ones receiving surveys (versus those at centers). This isn’t worded very clearly.**

Yes, only the teachers at awardee sites are receiving surveys. We have added additional text to the first sentence of the second paragraph to clarify our meaning. The sentence now reads:

As NASA centers are not required to recruit classroom teachers to participate in Sol, only teachers at national awardee sites will be administered surveys.

- 4. Page 6, on part B of the supporting statement, please explain why would a student be “unable” to complete the survey at the camp? This should be sufficiently defined to ensure even-ness of approach across awardee sites.**

We agree that providing opportunities for students to complete surveys outside of the official administration time could result in uneven survey administration conditions. After additional consideration of this issue, we have decided against providing additional opportunities for students to complete the survey and have removed this language from Page 6 (and an additional reference in Section B.3).

- 5. Page 7 and 10, supporting statement part B, We need to see a substantive plan to study nonresponse bias given that by the end of the each follow up as we are looking at a projected (cumulative) response rate of about 30%. NASA needs to figure out the details of this plan now because it may require collection of additional variables from camps, parents or students now to permit the analysis later. This plan should specific the methods and the variables to be used in the analysis.**

We thank you for highlighting the need for a substantive plan for nonresponse bias and have addressed your concern by adding an additional section to B.3:

Nonresponse Bias

Nonresponse may be a problem in our analyses if it introduces bias into our population estimates. Bias occurs if the students that refuse to participate or leave the study would give systematically different responses to the survey (had they responded to it) than the students who complete the surveys. Poor response rates do not guarantee a biased estimate, as the decision to not participate or leave the study could be completely unrelated to survey answers.

In general, the effects of potential non-response bias cause little concern if the non-response rate is less than 20 percent; accordingly, we will conduct a nonresponse bias analysis that is described below, if our response rate is less than 80 percent. We will construct a propensity model to estimate the probability of a student responding to the survey (propensity score) both for responding and nonresponding students. These propensity scores are estimated by a logistic regression model that will use demographic variables (e.g., gender, grade level, race, ethnicity) collected on the original parent consent form/survey that will be available both for nonresponding and responding students. We will then group students using the estimated propensity scores and examine the demographic characteristics of responding and nonresponding students within each group. This grouping will provide a method of forming weighting classes to adjust the weights of responding students and reduce nonresponse bias.

6. Page 13, supporting statement B, bullets 3 and 4 – please elaborate on the procedures to be used for the follow up – is this done at the camps? At home?

Below, please see additional information that we added for each bullet point about how we will ensure collection of this data:

- Asking teachers to fill out the baseline survey as a part of their registration packet. A webinar for national evaluation coordinators will be conducted prior to survey distribution that will emphasize the importance of collecting the baseline survey before the start of Sol programming. The national evaluation coordinators will be responsible for collecting these surveys for each camp.
- Giving teachers the opportunity to fill out follow-up surveys via an online link. Our procedures for the follow-up surveys include sending up to three email reminders and making up to three follow-up calls to encourage teachers to fill out the surveys at home or wherever they have internet access. Two national awardees have indicated that their teachers may not have internet access. For these awardees, we will print paper surveys and mail them to the national coordinator for administration.

7. Teacher baseline survey – the instruction page is way too long. NASA should only provide information needed to complete the survey since teachers should already have received information about the purpose of the study, etc. Therefore, we would suggest deleting the second paragraph, the first sentence of the third paragraph, and place the last two sentences of the last paragraph down with the PRA statement. Other trimming also should be considered.

We have pared down the instruction page to only include the information that our Internal Review Board (IRB) has deemed necessary.

- 8. Also on the Teacher survey, please array scales left to right not top to bottom. This is slightly less cognitively less burdensome.**

We have made these revisions to the layout of the scales.

- 9. Please provide a power analysis in Part B of the supporting statement. All the math starting on page 10 made us think we were getting to a power analysis, but we didn't seem to get there. We understand this is a formative evaluation, but if NASA plans to do a "pre-post" analysis, the Agency needs to be thinking about what you will have power to identify.**

We agree that it is important to think about what we will have the power to identify in our evaluation; see below for a description of our power calculations in Section B.2.1:

Power Calculations for Student Surveys

We have conducted power calculations to estimate the number of students that would have to be sampled at each awardee/Center to detect a minimum detectable effect size of .1 on science interest measures between survey waves with 80% power. This -assumes a two-sided test at 5% level of significance with a correlation between students across waves of .5 and a population standard deviation of 1.14. We consider a change of 0.1 to be substantive for purposes of deciding on changes for the project, as differences between the summer follow-up survey and the baseline survey ranged between -.1 and .1 in last year's pilot surveys. Based on results from last year's pilot, we assumed a response rate of 85% and an attrition rate of 30% for each follow-up wave. Because students are clustered within classes and classes are clustered within camps, the sample size was inflated using a design effect of 1.4 to account for intraclass correlation¹. The results of the power calculations, including adjustments for response rates, attrition, and the design effect, indicate that we will have to sample a total of 3,450 students from the national awardees and a total of 3,450 students from NASA centers. Thus, we will need a total sample of 6,900 students.

- 10. Please add some "control" questions to the student and teacher surveys, similar to what NASA will be doing on the NES evaluation. This allows for kind of a "difference in difference" estimate of the changes for something they expect Sol to change vs. something they do not expect Sol to change. This will give NASA some chance to account for students or teachers being systematically more positive or negative about science during the baseline or follow-up surveys. We are happy to chat more with NASA and Abt about appropriate "control" questions.**

We would appreciate a teleconference with OMB to discuss appropriate "control" questions. Our chief concern with adding these questions is that they would lengthen the survey which we are trying to keep at a burden of ten minutes or less; furthermore, it might prove difficult to anticipate all of the attitudes and behaviors that will be affected by participation in Sol as the activities range broadly.

¹ Note that the survey analysis team selected this design effect based on previous experiences with similar sampling designs and populations.

- 11. Like our passback on the NES package, for the student surveys a mix of positive/negative questions is preferable. Also, label all points on your scales – not just the 1st, 3rd, and 5th. We would also suggest moving the instruction to return the form to the teacher to the instructions for the child – not to include it in the PRA section which most of them will not read.**

As noted above, we have revised our surveys based on OMB's recommendation to intersperse positively and negatively worded questions. We adjusted the student surveys to include labels on all points of the scales. We agree with your concern that children will not read the instructions page of the survey in detail and have moved the instructions of what to do with the survey when they've completed it to appear right after the last question on the survey.

- 12. Appendix 12 – like the NES package, please get rid of the continue and no thanks option on the website.**

We have removed the no thanks option from the survey but have left the continue button as that is how we are moving respondents from the welcome page to the start of the implementation questions. The continue button is used to help the respondent progress through the form while saving their responses every few questions.

- 13. Appendix 14 – Do you want to ask any questions regarding NASA content more specifically? Are there certain modules which are never used?**

We have added a few questions to the focus group protocol to probe on sites' use of NASA content. We have also updated the implementation form to ask sites to indicate the specific content used and will assess the frequency of the modules that are used. This will allow us to identify which modules (if any) are never used and which are used most frequently. The data collected through the focus groups may help to provide a better understanding of why certain modules are used more than others.