**Panel Summary #1**

Intellectual Merit:

Several of the panelists loved the statement regarding the two major categories that students seem to fall into - those that innately love computer science and those who could be convinced given the right prompts. Recognizing that can lead to exactly the type of intervention that is desired within the IUSE program.

There was general agreement that while the big picture vision is correct; the mechanics are flawed. For example, can data analysis be done in blockly in an engaging way for students? There was significant discussion as to whether or not the topical nature of the data sets was appropriate for the audience as there was no detail of the 25 items, only three samples. Have there been any trials with students to test the methodologies and interest areas?

There was concurrence that while the project has merit, there may actually be a larger issue here. CS1 is currently not working with the current approach so let's try something else - augmented data sets. There needs to be retrospection to understand the core issues in the course and work to address those. It might be that this is the remedy, but it would require more evidence to be more convincing.

Broader Impacts:

The use of the OpenDSA is a positive and a strategic resource that can be leveraged. Getting the right remedy should be first and foremost and then it can be scaled within this environment.

There is a well thought-out evaluation plan, although there was concern that because of other possible underlying issues, measurement of the real impact of the effort could be clouded.

Summary:

The panel reviewed the proposal's data management plan and found it to be adequate.

There was agreement that this is a worthy endeavor and that addressing students who are not naturally inclined to study computing has merit. This proposal could be strengthened by addressing the concerns above.

**Review #1**

**Rating: Good**

**REVIEW:**

Short description of project:

This proposal seeks to create, deploy, and evaluate a revised emphasis in introductory CS at Virginia Tech, specifically to weave pedagogically rich datasets through a CT, CS1, and CS3 course. Student motivation, engagement, and self-regulation would be assessed in order to understand the effect of this "big data" approach.

Intellectual Merit:

Strengths:

First, the PI team, pedagogical support, and assessment effors on this project aroverall foundation are an experienced and well-established group: the human resources available to this project do inspire confidence in its progress, pacing, and success.

Branching from the expertise of the team is a second strength of this proposal - the OpenDSA framework is a (justifiably) NSF-supported pedagogical resource for scaffolding concepts in an accessible, engaging manner. This reader had not explored this resource before, but found it easy and enjoyable to use -- and certainly its incorporation here boosts confidence that the curriculum developed will be well-presented and widely and successfully disseminated.

This reader VERY much appreciated the vision cast on page 10, in which a single interface would provide both a blocks-based and a text-based description of the algorithm, with either one editable (and changes appearing in the other). This is already making substantial waves in the K-12 space through Google's Pencilcode project: it uses a different scripting language than this proposal (and is somewhat odd that it wasn't mentioned) -- but, on the whole, this does seem like the interface to which all computational communities (by any label) are converging for introductory students. This proposal's goals and plan benefit substantially from joining this path.

Concerns:

The use of Blockly for exploring large data sets did not seem natural or obvious to this reviewer - a much smaller pilot, e.g., at VaTech alone, suggesting efficacy might have helped alleviate those concerns. Similarly, the use of Java - a language designed for disciplined software-engineering - also seemed a mismatch for distilling insight from large, rich, multifaceted datasets, especially given that (at least in CS1), the language itself was presumably being taught. At the very least, the proposal could have made a stronger case against the vast majority of data analysts who use scripting languages, e.g., R, Matlab, Mathematica, et al., for their work (in favor of Blockly and Java).

Perhaps there is a larger tension here: a new, entirely exciting and worthwhile trend in computing (the early exploration of large, real datasets) is being shoehorned into toolsets designed for other purposes: Greenfoot, for the design-first approach to software development that undergird the department's introductory courses and Blockly, a language developed so that students would have a syntax-independent mental model of computational primitives (quite literally, the building blocks upon which future courses will rely).

To borrow from a conversation overheard at SIGCSE, this proposal rightfully chooses tools that support "CS for Software": robust and careful design-to-specification which seeks to express formally a set of processes already understood. But the novel activities proposed are of a very different ilk, namely "CS for Insight" -- the exploratory and poorly-specified computations which are tried and re-tried in order to develop deeper insights, rather than to express existing understanding. That is, the latter is throw-away in terms of software -- it's the insights that are important and that remain.

Perhaps, then, this proposal has reversed the emerging challenge for computing: it's bringing "CS for Insight" to other fields of study that now matters, not trying to forge software engineers out of people whose passions are for better understanding non-computational problems and their solution spaces, but would like to leverage the power of computing to help them do so.

This reader loved the emphasis on real datasets, carefully crafted to be "pedagogically rich" (a lovely turn of phrase), but the toolsets and the student activities did not sync well, at least in light of the past experience of this reviewer. The proposal echoed something similar - if not identical - in its past experience with NetLogo.

This reviewer was also a bit disappointed that the proposal did not list the 25 dataset ideas that it was proposing. It's possible they don't all exist yet, but proposals often can benefit by being iteration 1, intead of interation 0, on the path they are pursuing. The static analysis tools, similarly, seemed as though they might benefit from some smaller pilot runs before this larger effort: iteration 0 vs. 1.

In the context of the five review elements, pleaseevaluate the strengths and weaknesses of the proposal with respect to broader impacts.

Broader Impacts:

Strengths:

First, this proposal does an excellent job of describing its assessments, as well as the theoretical (and practical) frameworks from which they spring. This reader very much appreciated the foundation presented in Section 2: there seems little doubt that the creation of curriculum centered around meaningful, authentic artifacts and activities will hook those students, "who could be convinced to pursue computing as their major, if shown that the discpline provides a powerful tool for solving real-world problems." All science is computer science, indeed! [G. Johnson 3/25/01 NYTimes]

Concerns:

The annual high-school teacher workshop and Girls in Computing Day are both worthy efforts; this reader was not able to discern how or whether those events would be fed-back into the development of better, more accessible curriculum. It seems that the proposal seeks out audiences for compelling, accessible computing, but one suspects there may be such audiences already at Tech that might benefit from the close-knit, sustained collaborations that would be possible.

More generally, this project seems walled off by the name of the proposing department - specifically, one could imagine reaching over that nominal barrier to groups of students not well represented in computing and, even without proselytizing a different major course of study, at least providing resources that will make computing a tool that can help a \_huge\_ cohort of students deepen their abilities in their own fields. This reader is all for drawing students into computer science, but the true broader impacts are not within CS per se: they are in empowering students "who could be convinced to pursue computing as their major" to, instead, pursue their passions that need no coaxing, but, in doing so, to use computing to amplify their effectiveness and enjoyment of their work.

Please evaluate the strengths and weaknesses of the proposal with respect to any additional solicitation-specific review criteria, if applicable

[These have been incorporated in the above intellectual merit/broader impacts statements.]

Summary Statement

Overall summary and rationale:

In sum, this project's vision offers a good proposal for making introductory computing simultaneously compelling and accessible, and assessing the results within the department. For this reviewer, the computing focus is spot-on, but the departmental walls feel confining. There may be unexplored ways to further broaden the impact and audience for this proposal's beautiful pedagogical and content ideas -- so that they reach precisely the people who can benefit most from using computing to gain insight into non-computing questions.

**Reviewer 2**

**Rating: Good**

|  |
| --- |
| **REVIEW:**  *In the context of the five review elements, please evaluate the strengths and weaknesses of the proposal with respect to intellectual merit.*  This project would lead to the development of augmented datasets that can be used to engage  students with the topics of societal impacts and responsible professional behavior while  also exposing them to large datasets. The materials supplementing the datasets would be  designed to be usable in classes ranging from a CS0 computational thinking class to data structures. The plan to augment datasets is interesting. They will add layers enabling  students to look at both simple and more complex subsets of the data, describing the societal relevance of each dataset, giving research questions that students can explore using the data, and providing classification data to enable instructors to decide whether the materials are appropriate for their classes. It would be good to know whether this leads to higher rates of adoption.   Evaluation will be done using standard instruments. The materials will be tested in three classes: an intro to computational thinking class that is open to all majors, the first  programming class for CS majors, and a data structures class. The PIs estimate about 700  undergraduates will be affected. The evaluation would be stronger if there were also plans  to assess a control group.  The visualization component of the project does not seem as well integrated with the rest  of the proposal as it might be. More generally, they appeared tied to Greenfoot for reasons  that are not entirely clear. Given that they say their existing CS1 course causes "a large percentage of students [...] not just to withdraw, but to not pursue computer science as their major." it seems as if it might be worth looking into why that is. They hypothesize this is because of the game emphasis, but do not give evidence showing that this is, in fact, an issue that needs correction.   *In the context of the five review elements, please evaluate the strengths and weaknesses of the proposal with respect to broader impacts.*  Engaging students with issues related to ethics and societal impact in a meaningful way is  valuable. It can be difficult to get instructors to adopt new materials, but the design of  the materials in this proposal seems reasonable. Outreach to K-12 groups is also planned.   *Please evaluate the strengths and weaknesses of the proposal with respect to any additional solicitation-specific review criteria, if applicable*  Summary Statement   The idea of augmenting datasets with things that will make them more useful and usable is good. The proposal would have been stronger if it had contained more information about other aspects. |

**Reviewer 3**

**Rating: Good**

**REVIEW:**

*In the context of the five review elements, please evaluate the strengths and weaknesses of the proposal with respect to intellectual merit.*  
This proposal seeks to create and field test data-centric lessons in introductory programming classes through a robust learning platform on an Open EDX platform. The pedagogically-enriched resources will allow for scaffolded inquiry which is suggested will increase persistence in the affected courses and advancement in the major.   
  
Strengths: Data analysis is a critical skill today. Data is pervasive across all industries and having an understanding of how to approach data problems will increase competitiveness. The PIs have extensive experience in electronic book development and use in computer science education. The sample projects identified appear to have topical appeal.   
  
Weaknesses: The proposal seems to lack details about the students and relevant interests. There was a reference that a test offering concluded that the "inclusion of social impacts is motivating" without any details as to what and why.   
  
There was also mention of a large drop-out rate from one of the courses and subsequently from the discipline. I would have liked to have seen a multi-pronged approach for intervention, instead of suggesting this new methodology is the silver bullet.   
  
I am also not sure why the four characteristics (pedagogically-enriched data, social impacts, automatic feedback/visualization, and scaffolded environments are not present in all of the proposed courses. It seems that the proposal is built on the premise of the importance of these components.   
  
*In the context of the five review elements, please evaluate the strengths and weaknesses of the proposal with respect to broader impacts.*  
In addition to their own students, the PIs propose to engage middle and high school students as well. It will be critical to think about the different pedagogical needs for this audience and outline how this will be manifested.   
  
*Please evaluate the strengths and weaknesses of the proposal with respect to any additional solicitation-specific review criteria, if applicable*  
Summary Statement   
  
While there are many positive aspects of this proposal, there is little explanation about how the PIs and Co-PIs will manage the project and learn from the results. This needs to be addressed to make it more competitive.