Proposal Review 3 : 2321227

Back to Proposal

Agency Name:	National Science Foundation
Agency Tracking Number:	**2321227**
Organization:	
NSF Program:	CISE Core: Large Projects
PI/PD:	Pierce, Benjamin
Application Title:	CISE: Large: Property-Based Testing for the People
Rating:	Fair

Review

Summary

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to intellectual merit.

This proposal sits at the intersection of software testing and HCI. In particular, it focuses on property-based testing (PBT), an academic idea that has become increasingly popular in industry (and should become much more so!). It combines expertise in PBT with expertise in human factors work to examine and improve the ways PBT can spread. The tasks include needs-finding, better random generators, more kinds of PBT, interactivity, and dissemination.

***INTELLECTUAL MERIT

1. What is the potential for the proposed activity to advance knowledge and understanding within its own field or across different fields?

Very high. This is an important topic that has received very little academic attention other than purely mechanical issues (e.g., creating better generators, shrinking search spaces).

2. To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?

The tasks are both solid and interesting.

3. Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?

Pretty good, though I have several questions/comments below.

4. How well qualified is the individual, team, or institution to conduct the proposed activities?

Excellent.

5. Are there adequate resources available to the PI (either at the home institution or through collaborations) to carry out the proposed activities?

Yes.

Comments:

- I do not find JaneStreet a representative company at all. You might as well have surveyed an academic research group. I realize there is some text about mitigating this bias, but it is not persuasive.

- The proposd studies in 3.1 to generalize these studies do not address a key question of whether your recruiting processes will actually *generalize*. The kind of developer who attends these conferences, and attends a talk about PBT, is likely not qualitatively different from the kind of person you deal with at JaneStreet (even if they didn't go through the same kind of hiring process). In general, the proposal does not convince me the authors have thought through what it *means* to generalize, with the result that their recruiting processes may simply attract more of the "same old" (or at least "similar old").

- 3.2: Who are the subjects?

- 3.3: I would have liked to know more here. There are references to various other theories (information-foraging, Attention-Investment, etc.), but beyond the name-drops, there is not enough here about what *you* are doing.

- 4.5: I'm curious why Prof. Lampropoulos is not on this grant?

- 5.1: Why are "macros or meta-programming" "brittle"? They seem to be used quite effectively. It would have helped to know what kind of brittleness you have in mind, since you are arguing you don't have it.

- 5.2: To be honest, I really couldn't understand what this section is saying.

- 5.3: Model-based testing is ancient. I don't know why you think it hasn't been "well publicized". There is a ton of attention to it not

only in academia but also in industry. I worry that the PIs here are coming from too narrow a perspective to really know this area.

- 7.1: For what languages? Should we assume Python/Hypothesis?

- 7.3: How will this be measured?

--Strengths
- VERY clearly written (with very few exceptions)
- important topic
- nice interplay between formal topics and HCI
- variety of proposed deliverables

--Weaknesses
- background study unrepresentative
- what if the early needs-finding points elsewhere than the other proposed work? the parts feel disconnected
- doesn't seem to know there's a fair bit of attention to model-based testing
- several questions about details
- not really the kind of big, bold vision that fits the Large

Results of prior NSF support (if applicable):

Good.

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

1. What is the potential for the proposed activity to benefit society or advance desired societal outcomes (Broader Impacts)?

Several proposed, solid activities. The impact on industry and educational impact are clear. The mentoring and diversity could use more of a plan. The benefits to society are indirect but actually strong.

2. To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?

Somewhat. Nothing is especially creative or original, but I don't think that's a criticism.

3. Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?

Yes. I would have liked to see more of a plan for increasing diversity.

4. How well qualified is the individual, team, or institution to conduct the proposed activities?

Very.

5. Are there adequate resources available to the PI (either at the home institution or through collaborations) to carry out the proposed activities?

Yes.

--Strengths:
- several kinds of activities
- solid activities

--Weaknesses:
- diversity plan lacking

--Adequacy of Data Management Plan:

Adequate.

--Adequacy of Post-doctoral Mentoring Plan (if applicable):

N/A

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

-- Does the proposal identify a computer and information science and engineering grand challenge and an agenda to tackle such a challenge?

I don't think this rises to the level of a grand challenge. It feels pretty incremental.

-- Does the proposal explicitly identify the participating CISE core programs it covers and make the case for why the challenge is within the scope of one or more of these participating core programs?

Yes.

-- Does the proposal define the roles of all members of the team and the synergies among them in a Management and Coordination plan?

Yes.

BROADENING PARTICIPATION IN COMPUTING:

1. Goal and Context: Does the plan describe a goal and the data from your institution(s) or local community that justifies that goal?

Yes.

2. Intended population(s): Does the plan identify the characteristics of participants from an underrepresented group, including school level?

Yes.

3. Strategy: Does the plan describe activities that address the stated goal(s) and intended population(s)?

Yes.

4. Measurement: Is there a plan to measure the outcome(s) of the activities?

Yes.

5. PI Engagement: Is there a clear role for each PI and co-PI? Does the plan describe how the PI is prepared (or will prepare or collaborate) to do the proposed work?

Yes.

Overall, though, I find the BPC plan fairly perfunctory, especially the part about diversifying class TAs.

Summary Statement

Overall this is a solid proposal. As a pair of Smalls, this proposal
would be a no-brainer. I just do not see a grand challenge being
addressed for a Large. Combined with the several questions/concerns

about the work, I gave it a Fair.