Citizen Science for Rock Counting
Person X is used to show the separation between people's responses. The number does not necessarily correlate to a specific person's responses.

Question 1: What are the strengths of this approach to boulder mapping or rock counting?

Group A

Person 1

* Free labor
* Some amateurs are very good and motivated to do well (science is hobby)
* results may match experts if we get large response
* public engagement in mission
* inspire young scientists to do more
* many eyes better than few eyes
* may find things experts miss because they lack biases

Person 2

* public outreach potential
* reduces burden on science team members
* enhanced statistics (how many clickworkers will see each image?)

Person 3

* large amount of people allow for large amounts of data to be analyzed
* citizen science counts could be used as initial guesses or verification of other methods
* involves a large community of people enabling public outreach

Person 4

* results from citizen science can be as good as professional at much lower or zero cost and can possibly respond faster than scientists busy with spacecraft operations

Group B

Person 1

* more efficient than manual counting
* more accurate than automated counting or machine learning
* great public education/outreach benefits
* quantitative information available more immediately

Person 2

* available manpower may enable a "brute force" approach to problem solution

Group C

Person 1

* good statistics in principle quickly (??)
* humans able to deal with and flag problems

Person 2

* scalable ñ huge dataset could get crunched very fast in a pinch
* input of synthetic or calibrated data allows for weighting of users ñ good statistics
* public involvement; lots of it
* pre-built platforms with some of these groups (Zooniverse)

Person 3

* can analyze a lot of data very quickly
* quality of the data seems high
* engages a broad segment of the population in the mission
* provides multiple analyses of the same image
* brings a different perspective ñ may see something that we would miss

Person 4

* low cost, turn-key solutions to boulder counting
* good statistics, comparable to expert counters
* great for public relations
* addresses a major technical concern for the project without a large technical effort

Question 2: What are the weaknesses of this approach to boulder mapping or rock counting?

Group A

Person 1

* people are not paid, so response time may be long or limited
* outliers may dominate counts
* hard to trust result without experts checking things by doing it themselves
* must be well constrained problem that is suited for citizen science
* sizes, etc. will vary from user to user -> non-reproducible unless create instructions and rules

Person 2

* allowing the public means that you have untrained data analysts doing the work
* to ensure analyst techniques would need to provide some amount of training and validation

Person 3

* scatter of data is larger than otherwise
* scientific accuracy can be called into question
* needs time to determine methods, test site, write explanatory material
* no guarantee regarding when each analysis is complete

Person 4

* what if there is an interruption for some reason e.g. website going down at a critical time

Group B

Person 1

* may be conducted by amateurs so results may not be reliable

Person 2

* no security ñ images are released to the public completely pre-empting public relations value of mission accomplishments
* elaborate machinery to set up ñ needs tending by IT professional

Group C

Person 1

* concern for using unpaid volunteers gives you no leverage. You canít make many demands on volunteers. You need a backup position if your effort fails.
* CosmoQuest goes under, and you have to go to a backup plan

Person 2

* unpredictable output
* difficult to quality control
* another set of interfaces
* excitement may be difficult to pump up

Person 3

* because the trolls on high-profile task (unreadable) may be actively malicious workers, who can fool ìcat-(unreadable)î algorithms
* can we get the people when we need them, or will the fans have lost interest? Can we be sure of enough people at the time?
* Will the website overload/die at a critical time?
* Will it be taken seriously by NASA management/reviewers

Person 4

* Requires us to release data publicly early in the mission
* Need to make sure the project seems interesting and important ñ could come off as boring to

the public

Question 3: Are there any opportunities we can leverage to improve this approach?

Group A

Person 1
* choose problem well suited for public. Few arbitrary elements
* decide how to deal with outliers
* decide how much data will be analyzed by this approach
* decide how to treat results vs experts (who is right if disagreement)

Person 2

* provide online software guides for user community
* provide plenty of examples on what defines a boulder
* validate results with other techniques, especially manual counting
* use as initial guesses for other techniques that use a priori information

Person 3

* include counting from manual, image processing for some frames to validate
* rely on demonstrated successes from previous citizen science programs
* encourage inclusion in (upper level) classes to gather student interest to increase statistics and as formal outreach

Person 4

* maybe have a professional check the results of citizen science?

Group B

Person 1

* citizen science can be combined with all other approaches to improve both its quality and its efficiency; it can be combined with machine learning and automated techniques to improve efficiency; it implements manual counting so it already has an advantage in accuracy
* citizen science allows a more efficient program of experimentation and statistics-gathering to be implemented
* citizen science allows repetition of counting specific regions to be adopted; this potentially makes it the most accurate of all the techniques

Person 2

* (cannot read their response)
* use the manual counting as "training set" for supervised learning -> benefit of having a large training set

Group C

Person 1
* perhaps we can target our studies to science and engineering students and/or different age groups we are trying to target
* provide prizes for quality and quantity

Person 2

* use citizen science on low-important regions that overlaps with counting by other techniques

Person 3

* collaborations with other missions such as Dawn and Hayabusa 2

Person 4

* if this and automated match, suggests good quality. If not, suggests a systematic problem that may need attention

Question 4: What testing can we do to validate this approach?

Group A

Person 1

* lots of training cases, with problems highlighted
* try to reproduce results from asteroids previously explored

Person 2

* validate against an expert's manual counting for the same area
* use images from different times of date for the same exact spot to see if shadowing or sun angles affect ending results

Person 3

* compare to results from other counting techniques
* determine the minimum number of responses required per image for statistics to best match other methods

Person 4

* compare results on a "calibrated" data set that has been previously done and validated
* use a synthetic dataset with known characteristics

Group B

Person 1

* validation (how to validate) is a weakness in my opinion (relying on untrained people) -> however, testing a cross validation with (unreadable) techniques would be necessary -> (opportunity?)

Person 2

* counting of entire asteroid could be accomplished multiple times give enough participants allowing comparison and statistics
* compare citizen science results with manual and automated techniques

Group C

Person 1

* I think this technique has been validated against expert counters, although itís not clear that it has been used for boulder counting. If not is has to be validated against boulder/rocks.
* We could use the Sahr/Boynton data set
* Need to assess how data set could be protected against proprietary concerns

Person 2

* distribute real and artificial image, see whether success depends on that. Check how fast it gets done ñ how quickly do people start and finish

Person 3

* perform a comparative study with our own professional rock counters

Person 4
* do some small scale tests with real data ñ edge cases: smooth, rough, etc.
* real data and prizes may be needed to get interest for a trial
* need some way to test our expected response

Question 5: Are there other issues, questions, ideas that are not covered by the strengths, weaknesses, opportunities, and testing questions? (wildcard)

Group A

Person 1

* a trial run is a great plan, but we should try to find a best match dataset to what we can expect from Bennu to test or incorporate a range of potential test surfaces

Person 2

* making things into games or prizes to engage the younger scientists
* trial run is important for understanding what you need to provide the community to be successful

Person 3

* how much oversight is needed?
* how well posed are the problems?
* can we depend on this approach for short time issues, or must this be limited to long baseline times
* how difficult is it to set up website for analysis and keep things controlled

Person 4

* how do you insure or encourage participation during time critical periods, e.g. when critical responses are needed in short timescales

Group B

Person 1

* using social media to advertise the citizen science

Person 2

* great public outreach opportunity
* creates network of asteroid-savvy citizen scientists that can be tapped for future missions

Group C

Person 1

* major concern would be if there are proprietary concerns for the data that would concern the science team/NASA
* would need some guarantee of access to CosmoQuest or other software/technology if they go away (source code escrow agreement)

Person 2

* what is a rational timeline for getting results?
* how do we ensure we can maintain this as a needed capability
* how do we do quality control

Person 3

* how do we keep people interested on the day we need them?
* do we have "training". The most interested folks may burn out before we need them

Person 4

* data release. Issues with putting our data online early.