# Research Grant Reviewer Instructions

Report prepared by Harvard University Effective Altruism Student Group Philanthropy Advisory Fellowship

Developed on behalf of One Mind

Authors: Paul Oyler-Castrillo (team leader), Eric Gastfriend, Santhi Hariprasad, Gabrielle Molina, Kai Sandbrink

Last Updated: 1/12/2020

# Introduction

This document serves as a template of instructions to provide to scientific research grant reviewers. It covers the scientific literature on bias in science funding and best practices in grantmaking.

## Goals

The purpose of this project to have as great an impact as possible through the funding of scientific research on brain disease. We aim to spur innovations that will cut the disease burden from brain illnesses and injuries in half by 2050. In order to achieve this goal, we must evaluate research proposals not only on their scientific merits, but also on their strategic relevance to public health. We must attempt to move beyond incremental advances in knowledge and take risks to reach for transformational shifts in brain science paradigms.

## “People not Projects”

What if there was a better way to fund science, one that would lead to doubling the chance of finding a major discovery? It turns out, it’s been done before. Howard Hughes Medical Institute has an innovative approach that leads to great results:

> *“They fund science somewhat differently than NIH. They give longer grants, they give grants to people, rather than for a specific work and project, and so on. This paper from Pierre Azoulay looked at two populations that, by most observable characteristics seem pretty similar, and you look at those who received HHMI funding and those who did not. Azoulay concluded that HHMI grant recipients were 98% more likely to produce work that is in the top 1% by citations (Azoulay 2009)” (Bier 2019).*

Azoulay noted that HHMI “tolerates early failure, rewards long-term success, and gives its appointees great freedom to experiment,” whereas NIH recipients “are subject to short review cycles, pre-defined deliverables, and renewal policies unforgiving of failure” (Azoulay 2009). Another notable example of this approach to funding was British Petroleum’s Venture Research

Unit (VRU). According to a RAND analysis, VRU "aimed to fund determined researchers who questioned current thinking and would do transformative work. An important driver of the VRU approach was the idea that researchers with radical ideas would struggle to obtain funding through traditional means" (Lichten 2014). The results from just £20M in funding over 10 years were remarkable. "Work supported by the VRU led to creation of an enzyme company, Oxford Asymmetry, which sold for £316m in 2000, significant developments in green chemistry and several other notable achievements" (Lichten 2014). The program's director assessed that "of the 26 groups that were running at the initiative's close in 1990, perhaps 14 made transformative discoveries: that is, they did radically change the way we think, and several succeeded in achieving important scientific objectives that their peers had thought were impossible or irrelevant" (Lichten 2014). The MacArthur Fellows "Genius Grants" program also follows a similar "no strings attached" approach.

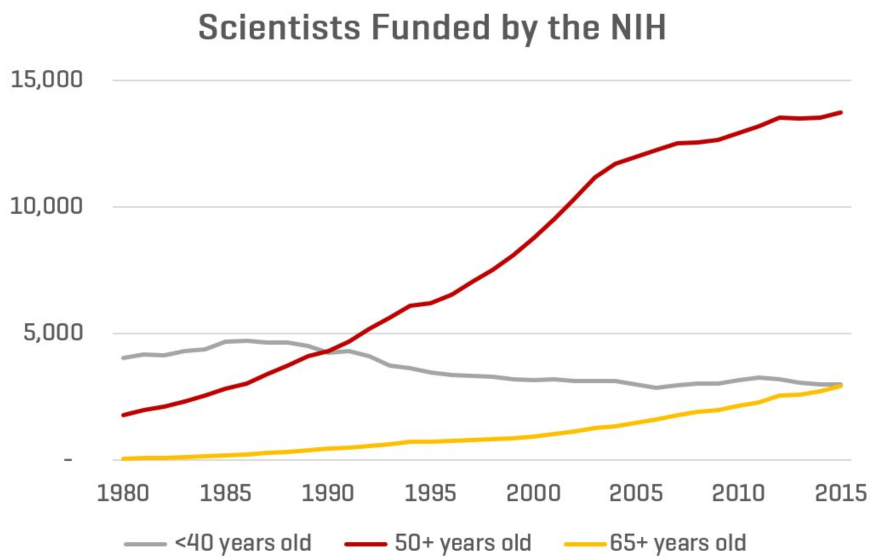
However, there are also some downsides to focusing too much on people instead of projects, such as hidden biases described in the following section. The key takeaway is to try to fund researchers and projects with big ideas that take risks.

# Hidden Biases

In surveying the scientometrics literature, we have found a number of systemic blind spots in science funding. Becoming aware of these blind spots is the first step towards correcting them:

## Researcher Age

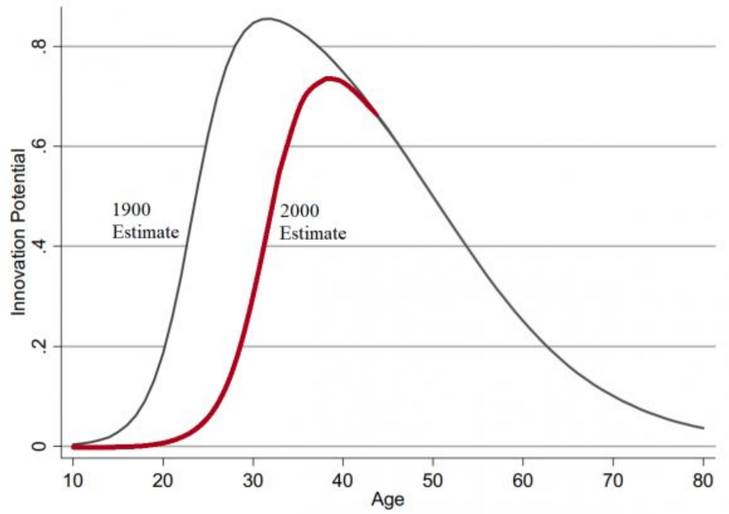
Funders may be overvaluing the achievement of older researchers and undervaluing the potential of younger researchers. Over the past 40 years, an increasing proportion of NIH funding has gone to PIs over age 50.




(Bier 2019)

This funding change may cause society to lose out on some of the most productive years of a researcher's career. Jones 2010 came to this conclusion by analyzing the likelihood of making great discoveries, such as those that lead to a Nobel prize, as a function of researcher age.

FIGURE 4.—MAXIMUM LIKELIHOOD ESTIMATES FOR THE POTENTIAL TO PRODUCE GREAT INNOVATIONS AS A FUNCTION OF AGE




The gap between the black curve (1900) and the red curve (2000) shows the lost potential for innovation at younger ages, due to researchers starting to do research later in their life (Jones 2010).

## Team Size

Research team size has increased dramatically over the past 40 years.

**Average Number of Authors per Paper**

| | 1980 | 2016 | increase in coauthors |
|---|---|---|---|
| Biology | 1.8 | 4.9 | +178% |
| Biomedical Research | 2.6 | 7.0 | +167% |
| Botany | 2.1 | 5.1 | +142% |
| Cancer | 3.6 | 8.6 | +139% |
| Internal Medicine | 2.7 | 6.3 | +131% |
| Physics | 2.6 | 5.8 | +123% |
| Chemistry | 2.8 | 5.2 | +87% |

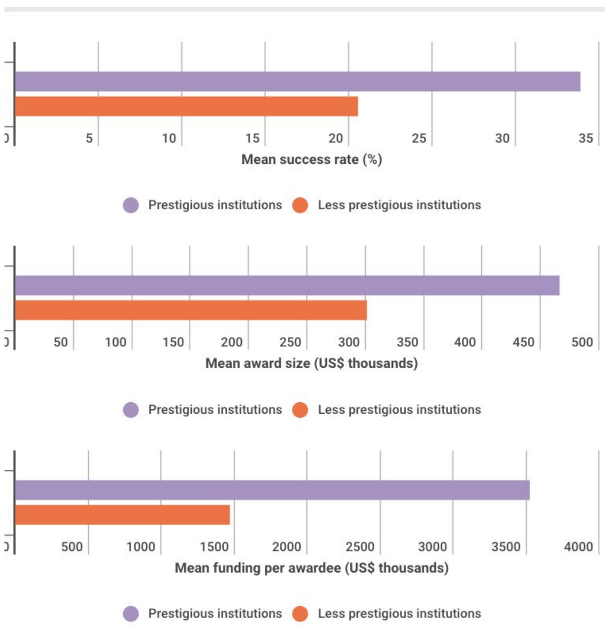
(Bier 2019)

However, an analysis of 65M papers and patents found that “smaller teams have tended to disrupt science and technology with new ideas and opportunities, whereas larger teams have tended to develop existing ones” (Wu et al 2019).

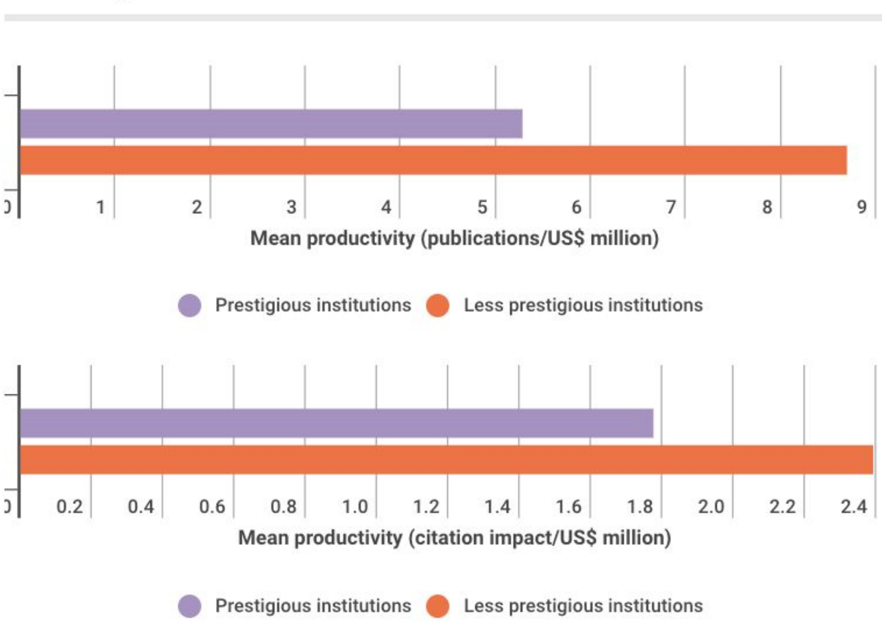
## Prestige

It should come as no surprise that researchers from elite institutions are more likely to get their NIH grants funded.

Analysis of 15 institutions finds NIH favours elite institutions.



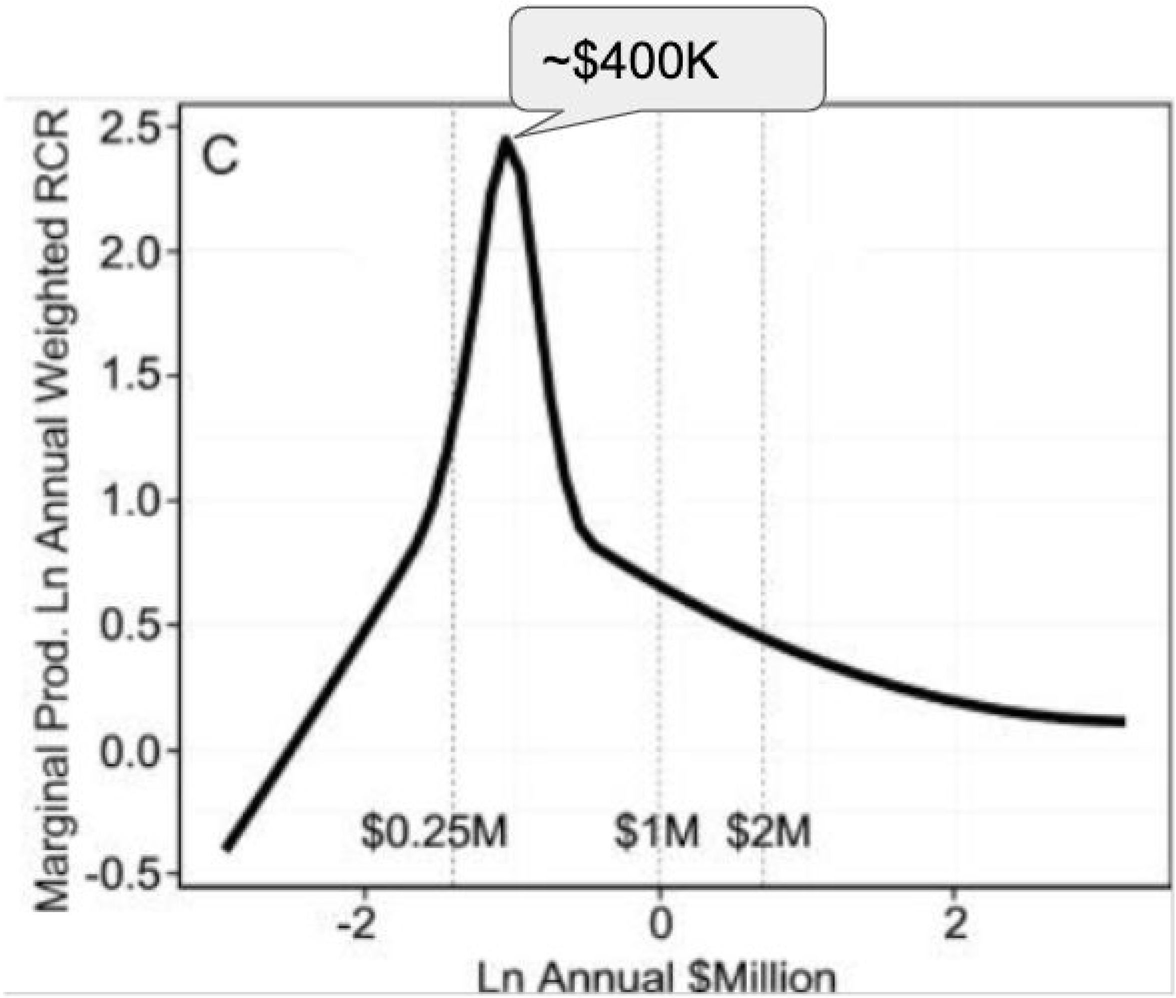

However, it may come as a surprise that by objective metrics, such as publications and citation impact, grants go further in less favored institutions.




Source: Wahls 2018 data. Catherine Armitage, "Less prestigious institutions deliver better value for grant money," Nature Index. Oct 18 2018.

## Award Size

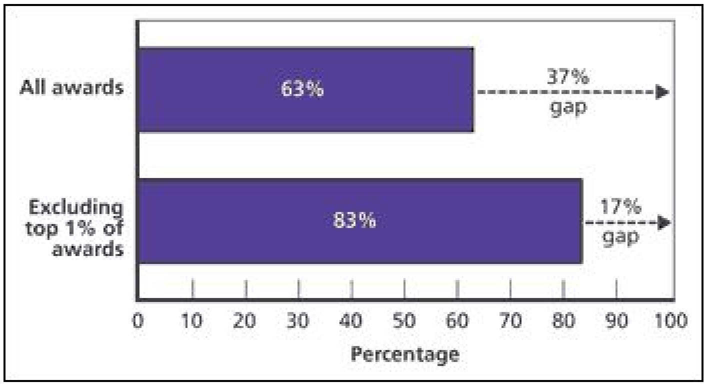
Scientometrics has also shed light on the optimal award size for grants. Although this of course depends on the nature of the project, a broad-strokes analysis found that productivity (in terms of citations produced) peaks at approx. $400K per investigator.




Relative Citation Ratio (RCR) is a measure of influence developed by the NIH, looking at the number of citations, adjusted for field of study and year of publication. (Wahls 2018)

## Gender

Although NSF and USDA research grants have parity between male and female researcher applicants, NIH funding has been found to have a significant gender gap. “As shown in the top bar, female applicants in FY 2001–2003 received on average only 63 percent of the funding that male applicants received (a gender gap of 37 percent).” (Hosek 2005)

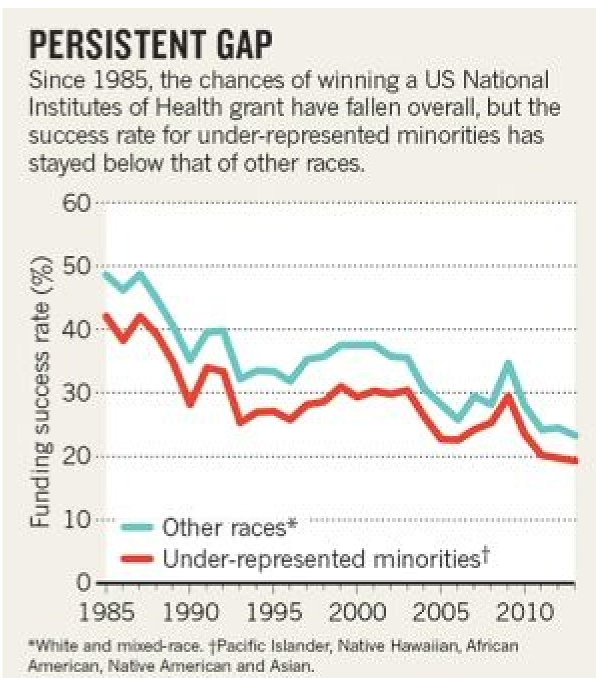



Another study found that among first-time NIH grant recipients, women are awarded 24% less than men, even after controlling for multiple confounders (Uzzi 2019).

An empirical study from Canada found that the gender gap among grantees may be due to reviewers' assessments of CVs: "women are less successful in receiving funding if reviewers are explicitly asked to review the principal investigator, rather than when they are asked to assess the quality of the science." (Witteman 2019).

## Race

NIH grants have also been found to have a racial disparity:




(Hayden 2015)

The drivers of this gap are subject to much debate. According to the authors of the most comprehensive analysis on the subject (Ginther 2018): "While we indeed considered bias as a possible explanation for some of our findings, we emphasized that bias, if present, was not likely to be the primary driver of the differences in funding that we found" (Kington 2018). They concluded that addressing the gaps in the training process for minority researchers may be the best way to shrink the gap. However, observed variables still do not fully explain the gap, leaving room for reviewer bias. The authors write: “It is unreasonable, however, given the evidence of the pervasiveness of bias throughout numerous dimensions of life, that scientists are exempt from our tendency toward bias, and equally untenable to suggest that scientists on NIH review committees are somehow the exception and exhibit no bias in their review of grants or, more importantly, in their day-to-day activities as working scientists” (Kington 2018).

# Scorecard Instructions

## Point System

For simplicity, we use a 0-10 scoring system, with 0 being the worst and 10 being the best. Please note that this is different from the NIH scoring system.

In order to ensure fairness, we ensure that every reviewer maintain 5 as the average score. In order to achieve this, we ask that you allocate yourself a pre-set number of points, based on the number of applications you're reviewing, to be divided among the applications.

Total Points Allocated = 5 * [Number of Applications] * [Number of Questions per Application]

If your average score across applications is not 5, we will re-normalize your scoring to a mean of 5 to ensure that reviewers' evaluations are comparable.

We recommend reading through a few applications to get a sense of the average quality before you begin scoring.

## Investigator(s)

In this section, you will score the Investigator based on question 1. If there are two co-applicants, you will score each and take an average of the scores. The investigator scoring falls into two categories, each weighted 50%:

### Accomplishment relative to experience

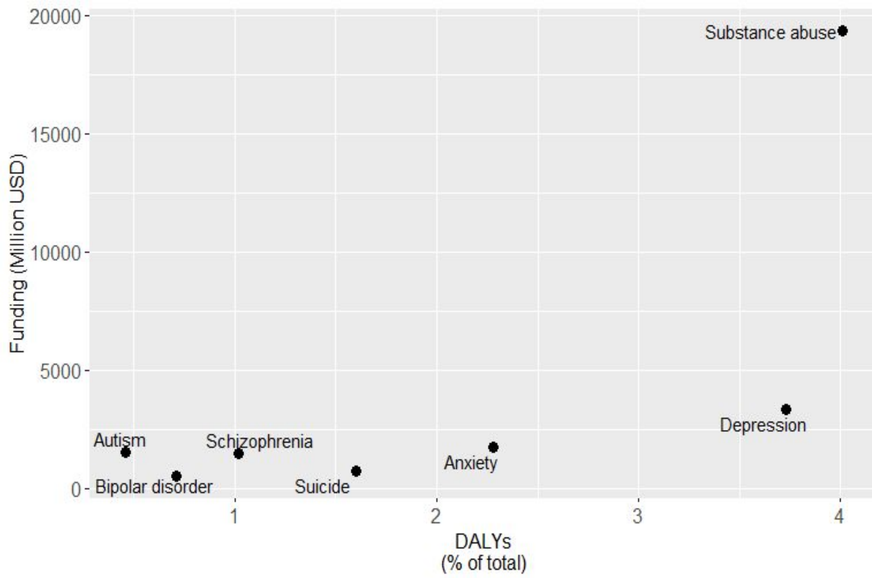
Score the applicant's past accomplishments on a scale from 0-10, taking into account the number of years since their first professional position post-training. You can consider accomplishments outside academia, such as companies founded or patents.

### Expertise in Research Area

Score the applicant's expertise in the research area on a scale from 0-10 based on publications, teaching, and positions held.

## Disease Area Neglectedness

In an ideal world, we would expect research funding to be allocated to diseases proportionately to how much suffering each disease imposes on society. Unfortunately, due to numerous social and political factors, this is not the case. There are "popular" and "unpopular" disease areas. We can have an outsized impact if we allocate more funding to the "unpopular" diseases, especially if they are overlooked not because they only affect small numbers of people, but because the government or society has neglected a large problem.




We measure neglectedness by looking at the amount of funding that the NIH (the primary funder for mental health research) allocates to different disease areas. We then compare that to importance, which we measure using Disability Adjusted Life Years (DALYs), a measure which combines public health burden from both morbidity and mortality. According to this analysis, substance abuse and autism are the least overlooked disease areas, and suicide is the most overlooked.

## Cost-Effectiveness of Research

A simple cost-effectiveness estimate can help maximize impact. *This calculator provides a rough approximation of cost-effectiveness, solely for the purpose of comparing research proposals*. There are two relevant questions in this calculation: (1) what is the potential impact of this research and (2) how much will this research cost?

1) Impact has three components:

- ***Disease Burden--DALY rate of condition in the United States:*** Like in the last category, public health burden is measured using DALYs.
- ***Reach--percentage of population reached by intervention:*** Some interventions are only applicable to a subset of the population affected by a condition (e.g. individuals who have been recently diagnosed). Even interventions that are applicable to many patients are only adopted by some of them (e.g. antidpressants).
- ***Effectiveness--average reduction in DALYs per individual reached:*** Successful interventions may reduce the prevalence of a health condition by preventing or curing cases. Alternatively, an intervention may reduce the length or severity of an illness, or prolong life.

2) Cost of Study: Some research is more expensive to conduct. One can have a greater impact with its funds by supporting research that is relatively low-cost.

Once you've calculated the cost-effectiveness of all the research proposals, the research proposal with the highest cost-effectiveness will receive a 10. To assign a score to the rest of the proposals, divide the cost-effectiveness score of the top-rated proposal by the one you are currently assigning a score to, then multiply by 10. For example, if the top-rated proposal had a cost-effectiveness of $25/DALY, the score for a proposal with a cost-effectiveness score of $40/DALY would be (25/40) *10 = 6.25.

# Scientific and Technical Merit

This is the section that most closely resembles the NIH funding application. This section is weighted at 50% for basic science research proposals, and 20% for all other proposals.

## Significance

On a scale from 0-10, how **significant is the potential scientific impact** of this work? Your score should be based on the applicant's answer to question 3 as well as your own analysis of this project. Keep in mind the following questions from the NIH's scoring system:

> "Does the project address an important problem or a critical barrier to progress in the field? If the aims of the project are achieved, how much will scientific knowledge, technical capability, and/or clinical practice be improved? How much will successful completion of the aims change the concepts, methods, technologies, treatments, services, or preventative interventions that drive this field?"

## Methodology/Approach

On a scale from 0-10, **how appropriate is the study's methodology**? Your score should be based on the applicant's answer to question 3 as well as your own analysis of this project. Keep in mind the following questions from the NIH's scoring system:

> "Are the overall strategy, methodology, and analyses well-reasoned and appropriate to accomplish the specific aims of the project? Have the investigators presented strategies to ensure a robust, appropriate, and unbiased approach? Are potential problems, alternative strategies, and benchmarks for success presented? If the project is in the early stages of development, will the strategy establish feasibility and will particularly risky aspects be managed?"

## Innovation

On a scale from 0-10, **how innovative is this proposal**? Your score should be based on the applicant's answer to question 3 as well as your own analysis of this project. Keep in mind the following questions from the NIH's scoring system:

> "Does the application challenge and seek to shift current research or clinical practice paradigms by utilizing novel theoretical concepts, approaches or methodologies, instrumentation, or interventions?"