U.S. Department of Health & Human Services

NIH Center for Scientific Review

CSR Update Council of Councils

May 2017

Richard Nakamura, Ph.D., Director
Center for Scientific Review

CSR Mission

To see that NIH grant applications receive fair, independent, expert, and timely reviews – free from inappropriate influences – so NIH can fund the most promising research.

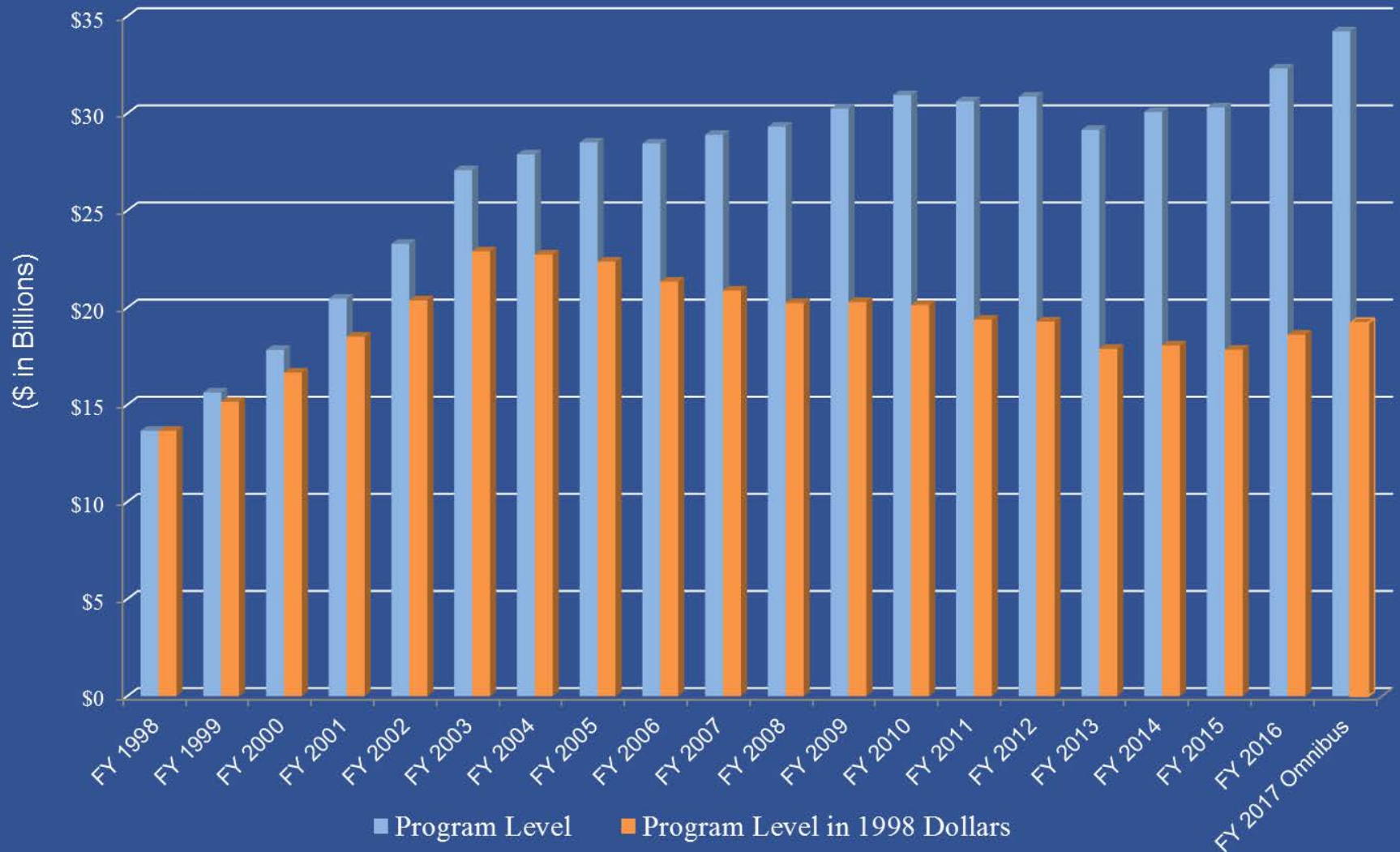# CSR Peer Review – Fiscal Year 2016

- 92,000 applications received

- 60,000 applications reviewed

- 18,000 reviewers

- 247 Scientific Review Officers

- 1,600 review meetings
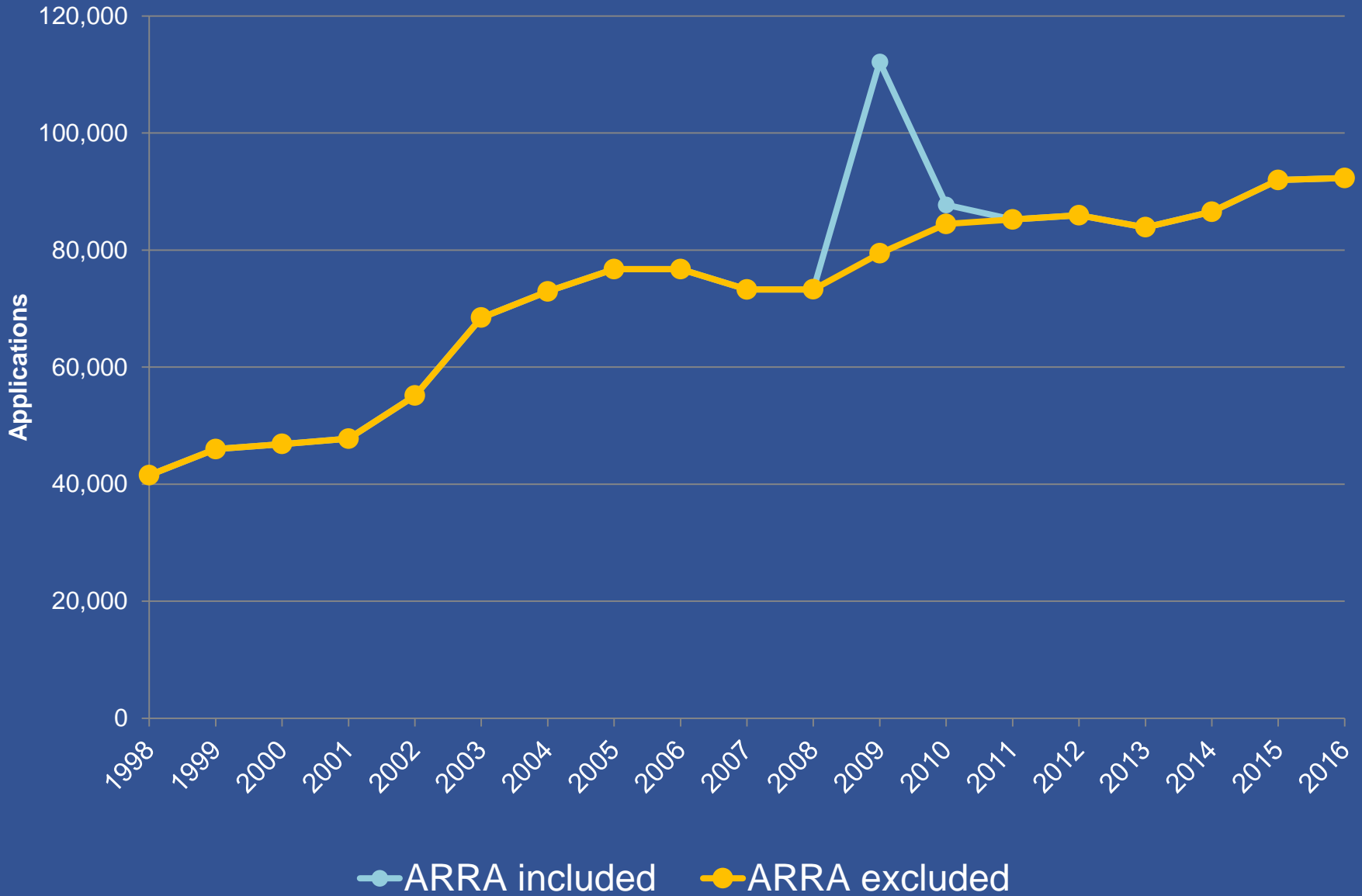
Issues for Science and for CSR

- Stagnant funding, many applications, low success rate
- Reviewer fire, reviewer burn out
- Survival of an excellent system of science
  - Signs of trouble
  - Rigor and reproducibility
- Measuring quality so the system can be improved
  - Bibliometric and scientist based measures
- Bias in review – study of anonymization
- Some other studies of review
  - Application distribution, half point pilot, grant ranking
  - Collaborative study with NIGMS
- Improving the process of review and award
- Good news

NIH Program Level
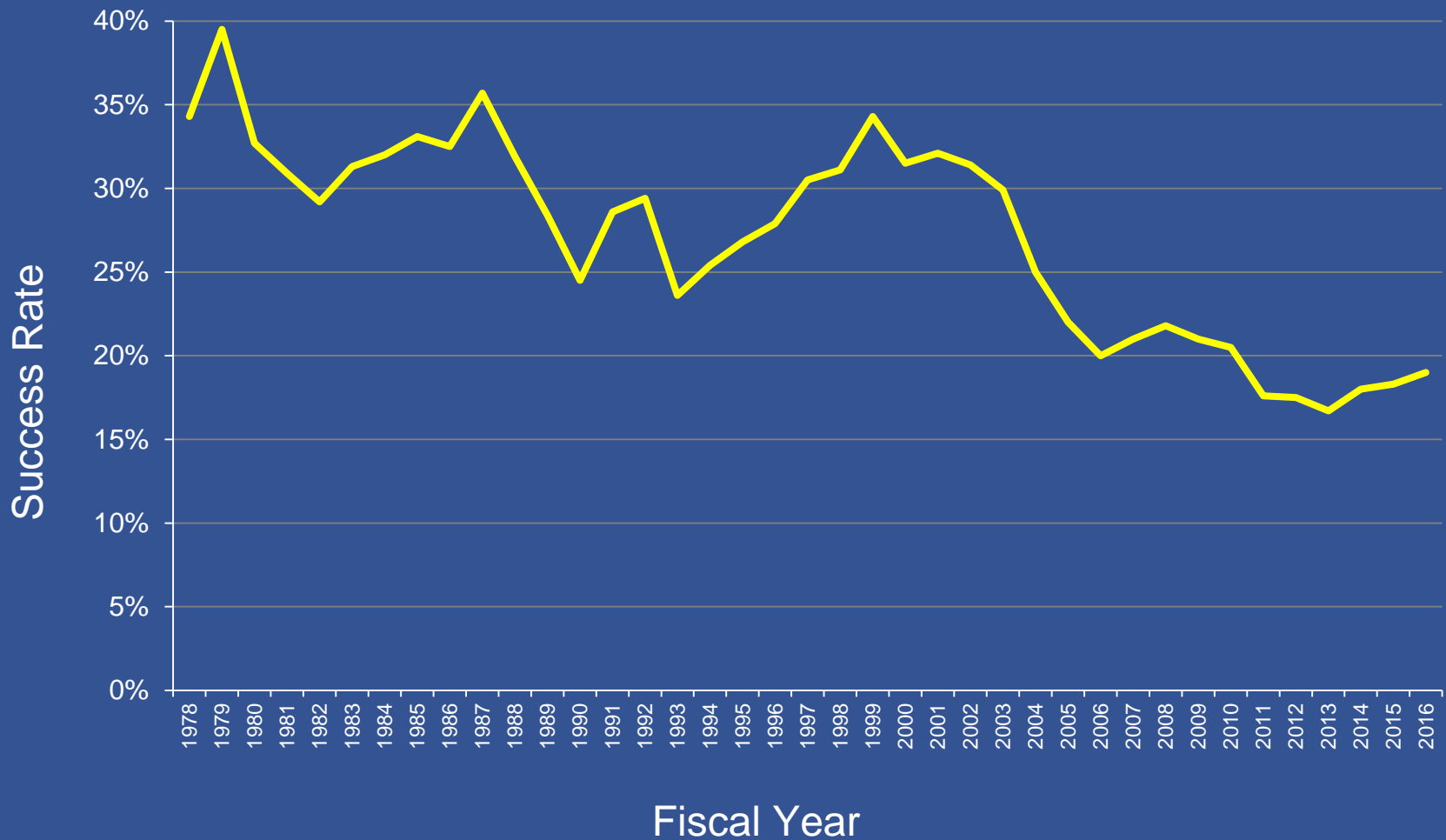## in Nominal Dollars and Constant Dollars

($ in Billions)

FY 1998, FY 1999, FY 2000, FY 2001, FY 2002, FY 2003, FY 2004, FY 2005, FY 2006, FY 2007, FY 2008, FY 2009, FY 2010, FY 2011, FY 2012, FY 2013, FY 2014, FY 2015, FY 2016, FY 2017 Omnibus

■ Program Level   ■ Program Level in 1998 Dollars

Number of Applications Received by Fiscal Year

Grant Success Rates
FY 1978 – 2016

U.S. Department of Health & Human Services

NIH Center for Scientific Review

Rigor and Reproducibility

Role of Peer Review

Reviewing Rigor and Transparency of Research:
RPG Applications

	Applies to which applications?	Where will I find it in the application?	Where do I include it in my critique?	Addition to review criteria	Affect overall impact score?
Scientific Premise	All	Research Strategy (Significance)	Significance	Is there a strong scientific premise for the project?	Yes
Scientific Rigor	All	Research Strategy (Approach)	Approach	Are there strategies to ensure a robust and unbiased approach?	Yes
Consideration of Relevant Biological Variables, Such as Sex	Projects with vertebrate animals and/or human subjects	Research Strategy (Approach)	Approach	Are adequate plans to address relevant biological variables, such as sex, included for studies in vertebrate animals or human subjects?	Yes
Authentication of Key Biological and/or Chemical Resources	Project involving key biological and/or chemical resources	New Attachment	Additional review considerations	Comment on plans for identifying and ensuring validity of resources.	No

Scientific Premise:
Guidance for Reviewers

GOAL: Ensure that the underlying **scientific foundation** of the project—concepts, previous work, and data (when relevant)—is sound.

- Pertains to the strength of the scientific foundation of **evidence/data** that increase possibility of high impact for the project

Premise should not be confused with hypothesis or significance

Scientific Rigor:
Guidance for Reviewers

GOAL: Ensure a strict application of scientific method that supports robust and unbiased design, analysis, interpretation, and reporting of results, and sufficient information for the study to be assessed and reproduced. Give careful consideration to the methods and issues that matter in your field.

NIH Center for Scientific Review

Sex as a Biological Variable:
Guidance for Reviewers

NIH expectations for **reviewers**:

- As part of the Consideration of Relevant Biological Variables, assess whether the plans to address sex as a biological variable are adequate (for studies in vertebrate animals or human subjects).

- If the study involves only one sex, is this justified scientifically?

- Assess within the context of the research question and current scientific knowledge.

Plan for Resource Authentication:
Guidance for Reviewers

GOAL: Ensure processes are in place to identify and regularly validate key resources used in their research and avoid unreliable research as a result of misidentified or contaminated resources.

- Researchers are expected to authenticate key biological and/or chemical resources used in their research, to ensure that the resources are genuine.
- New Review Consideration
- Reviewers rate as acceptable/unacceptable (provide brief explanation if unacceptable)
- Does not affect criterion scores or overall impact score

Related review issues:

- Different research fields may have different best practices for and reach different conclusions about scientific premise and rigor. Use the words. Assess based on best practices in the field.

- Page limits have not changed

- Cost of larger subject populations

- Good science can emerge from different styles

- More background investigation of premise

- The reviewer is the judge of premise and rigor

- Exploratory studies are still allowed

- Significance and potential for impact are still important
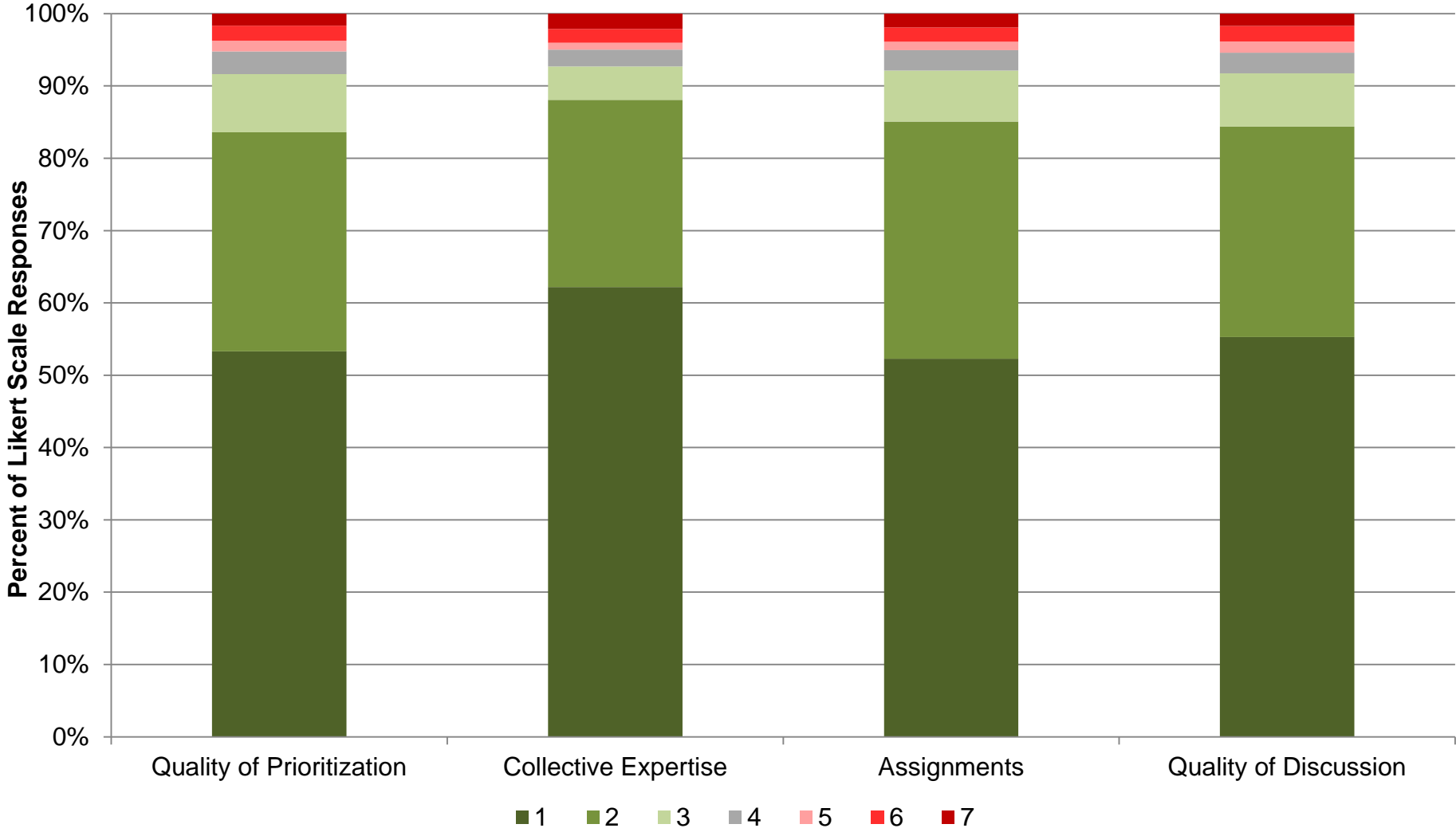
Some CSR Studies

What is CSR doing to improve review quality?

We are emphasizing evaluations by our scientific community and evolutionary improvements of staff, organization, and processes.

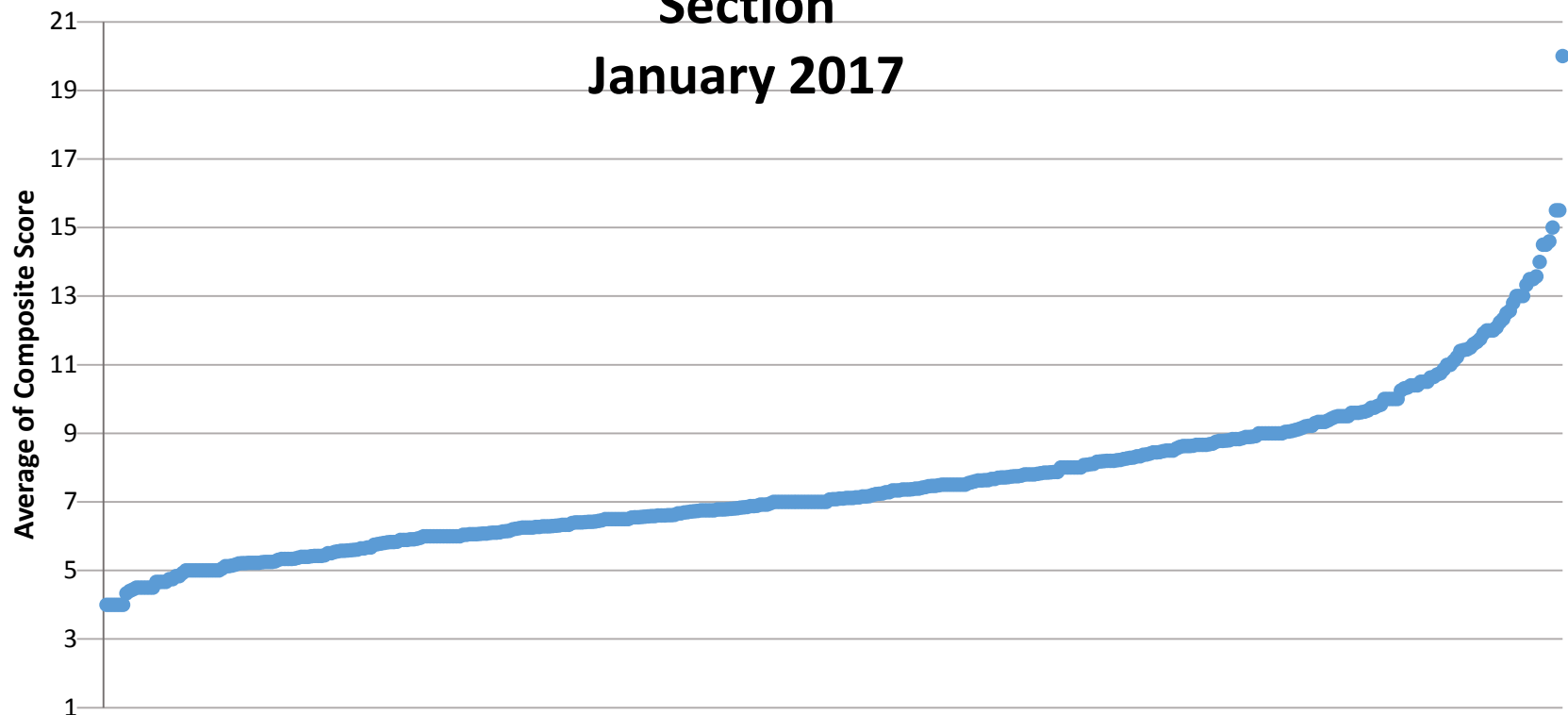# CSR Survey of Reviewers - January 2016

Likert type scale where 1=very strongly agree to 7=very strongly disagree. Overall NIH response rate = 50% (n=7,094), 267

NIH › Center for Scientific Review

Average of Reviewer*Composite Scores by Study Section
January 2017

Lowest = 4
ZRG1 AARR-K (50), ZRG1 GGG-E (02)
ZRG1 AARR-M (54), ZRG1 HDM-A

Median = 7.08
ZRG1 BST-W (80)
ZRG1 OTC-B (55)

Highest = 20
ZRG1 IMST-L (02)

Overall n = 10,235, Response rate = 47%, 507 minus 65 CSR study sections with fewer than 10 reviewers (n=442) except for Telephone Assisted Meetings.

CSR Conducts Regular Surveys and is Responsive to Input and Suggestions
from NIH and IC Leadership

CSR Program Officer Survey 2015 vs. 2016

Likert-type scale where 1 = strongly agree to 7 = strongly disagree. 2015 overall n = 916, Response rate = 38%. 2016 overall n = 905, Response rate = 37%.
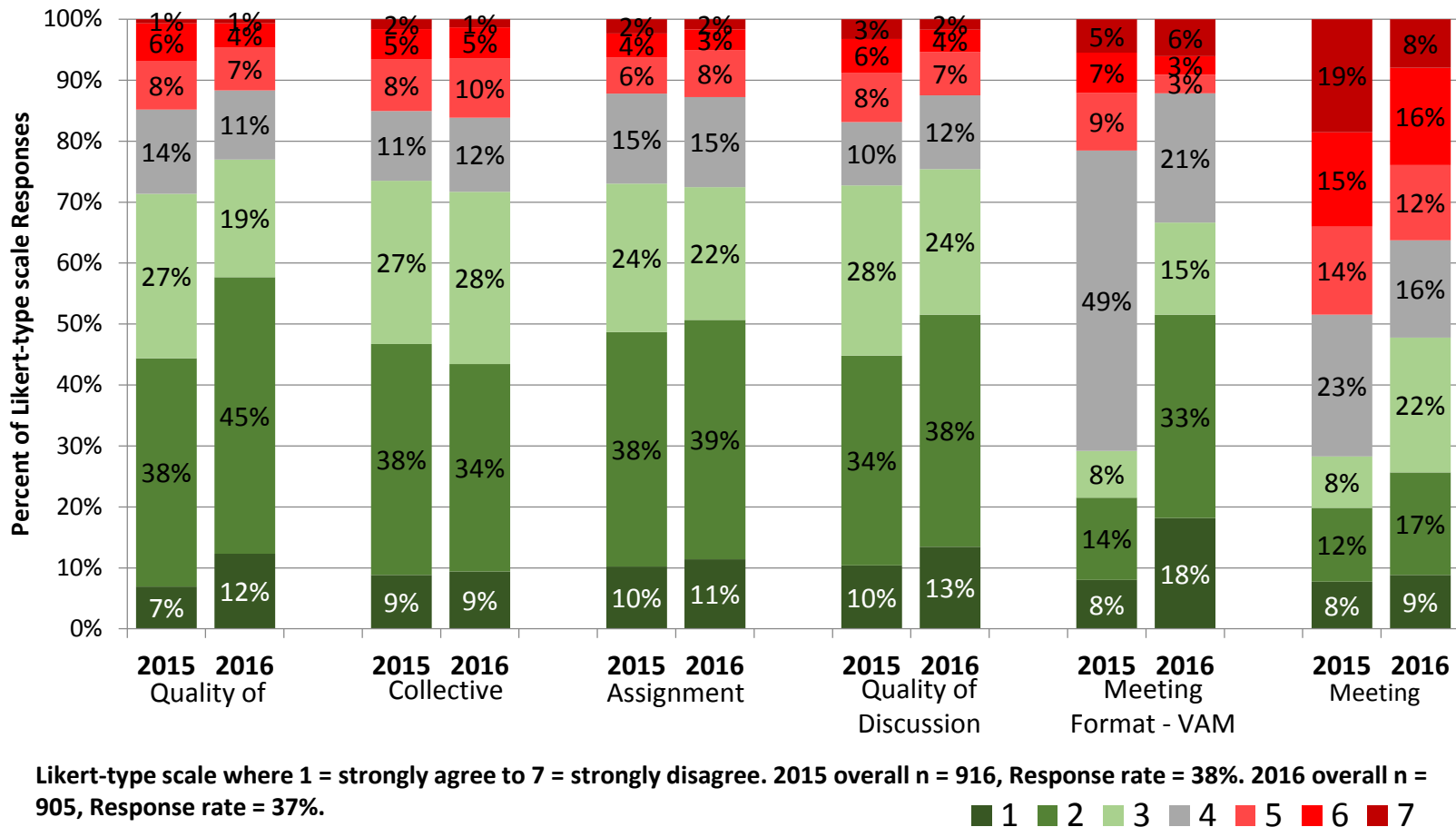
U.S. Department of Health & Human Services

Center for Scientific Review

NIH

Final Scoring of Applications in Half Point Increments: Pilot Study

Amy Rubinstein, Ph.D.
Scientific Review Officer

Current System for Evaluating and Ranking Applications Reviewed in CSR Study Sections

- Applications are assigned to 3 reviewers who provide preliminary impact scores (1-9) and critiques.

- After panel discussion of each of the top 50% of applications, all panel members vote on a final overall impact score in light of the range of scores supplied by the assigned reviewers.

- Each application's score is derived from the average of all panel members' votes and multiplied by 10 (resulting in final scores of 10-90).

- R01 applications are assigned a percentile based on the scores of applications reviewed in the relevant study section in that round and the previous 2 rounds.

Why Consider Adding Half Point Increments?

- Concerns about funding results in compression of scores in the high impact 1-3 range (final scores between 10 and 30). This range offers few choices in quality judgements.

- Tied scores, particularly at 20 and 30, are frequently encountered.

- Score compression and tied scores can make funding decisions challenging.

- The ability for reviewers to score applications with half point increments may help spread scores and allow reviewers' scores to more precisely reflect their true ordering of the likely impact of applications in a study section.

Distribution of Application Scores
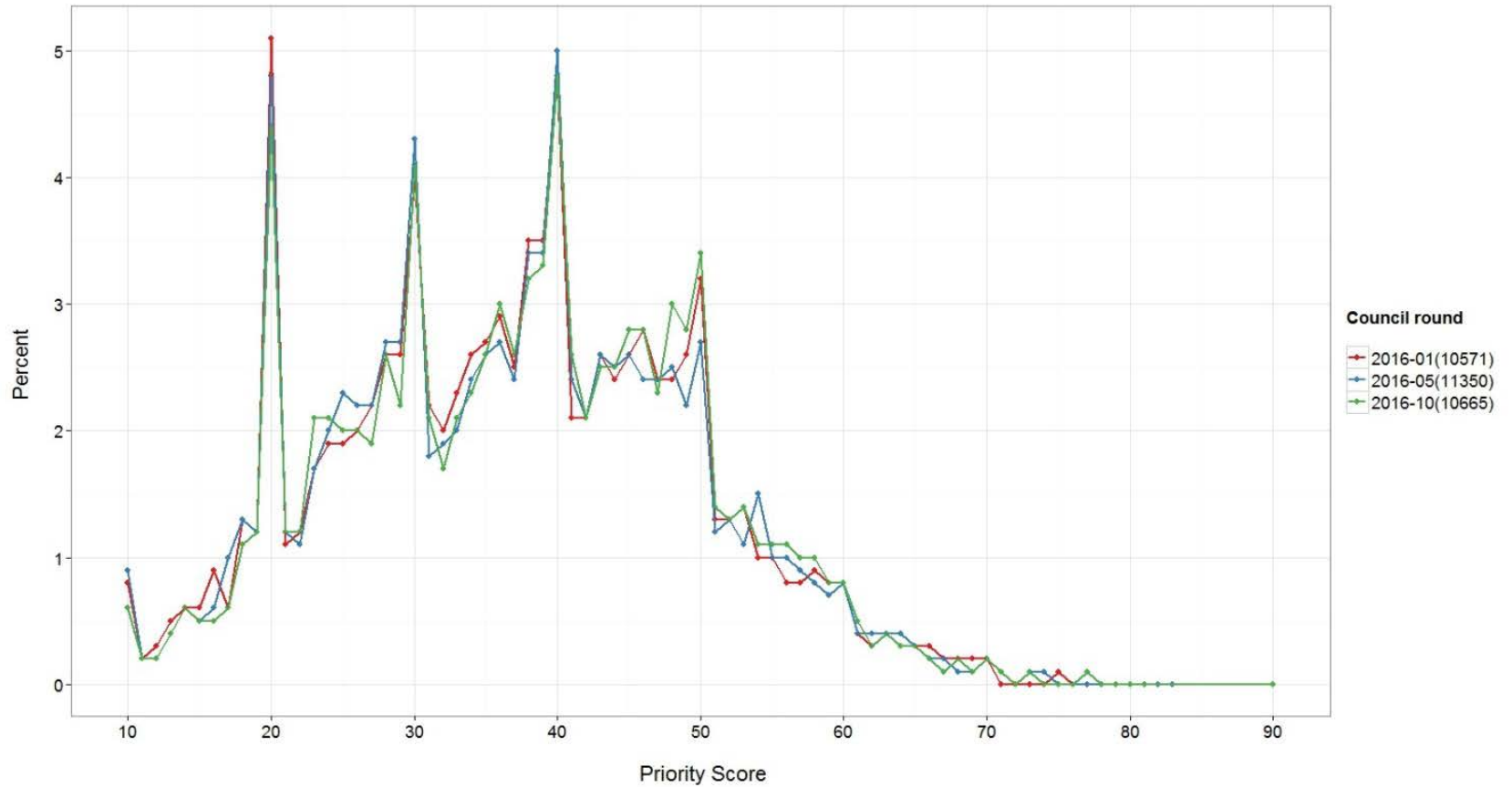
Half Point Pilot Study

- Carried out in parallel with the current review system in the 2016-05 and 2016-10 Council rounds.

- Applications were scored as usual to generate official scores for the applications.

- In final voting after the discussion of each application, reviewers were invited to provide a second score for each application that could range from a half point better to a half point worse than their official score.

- Half point scores were used to calculate alternate scores for each application to compare to the official scores. All analyses were produced for internal purposes only and alternate scores were not communicated to either applicants or program officials.

# How Did the Half Point Option Affect the Distribution of Application Scores?

1371 applications

in 39 study sections
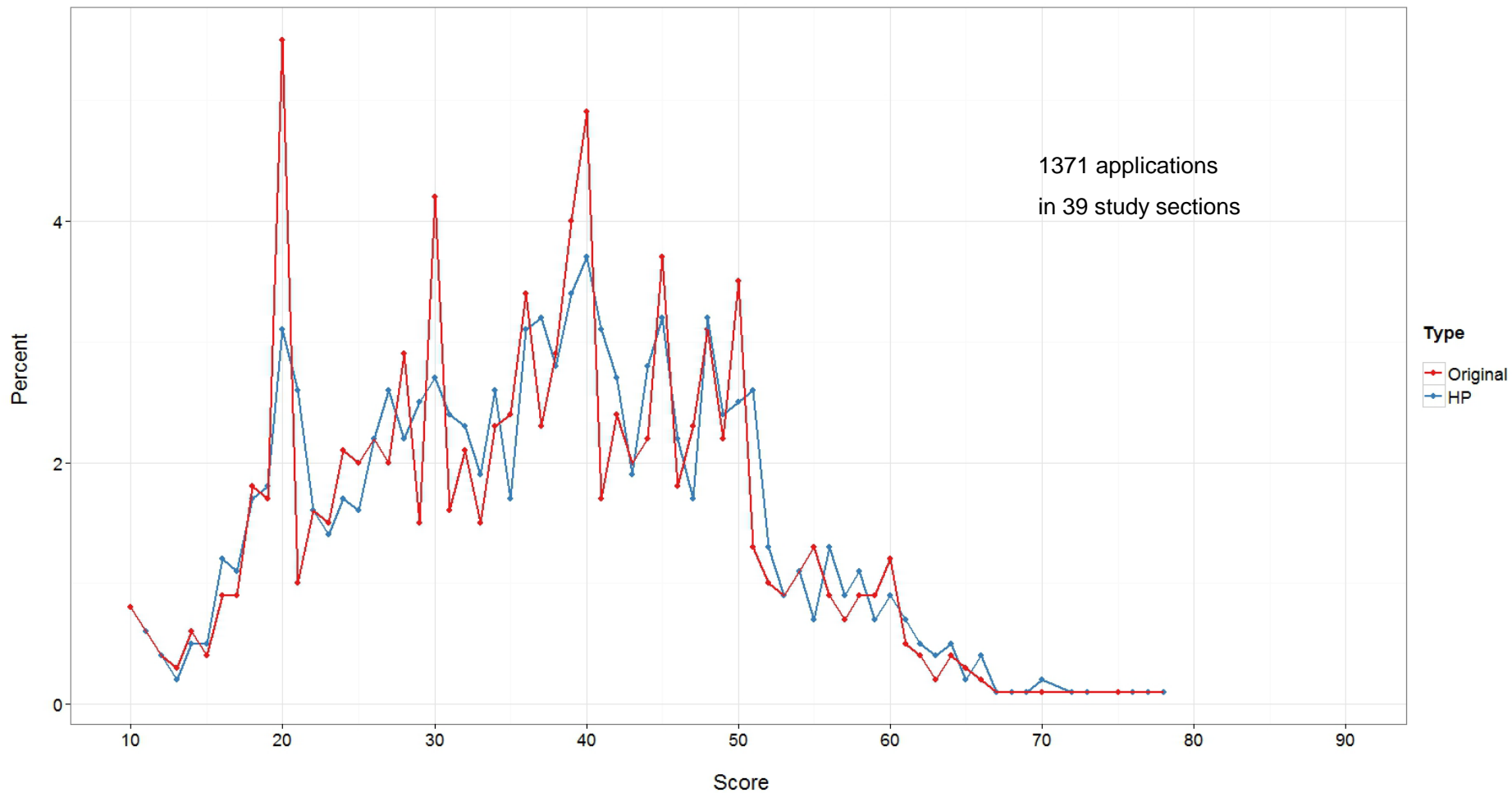
Survey Results: Reviewer Usage of Half Point Option in the 2016 10 Council Round

14%

86%

■ Used option

■ Did not use option
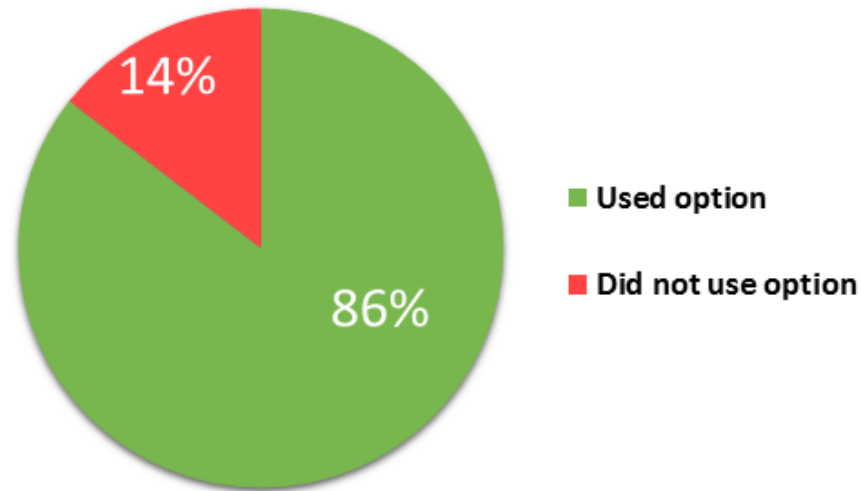
Overall response rate = 44%, n = 311.

How Did Reviewers Feel About the Half Point Option?

Likert type scale where 1 = Strongly agree to 7 = Strongly disagree. Overall response rate = 44%, n = 311.
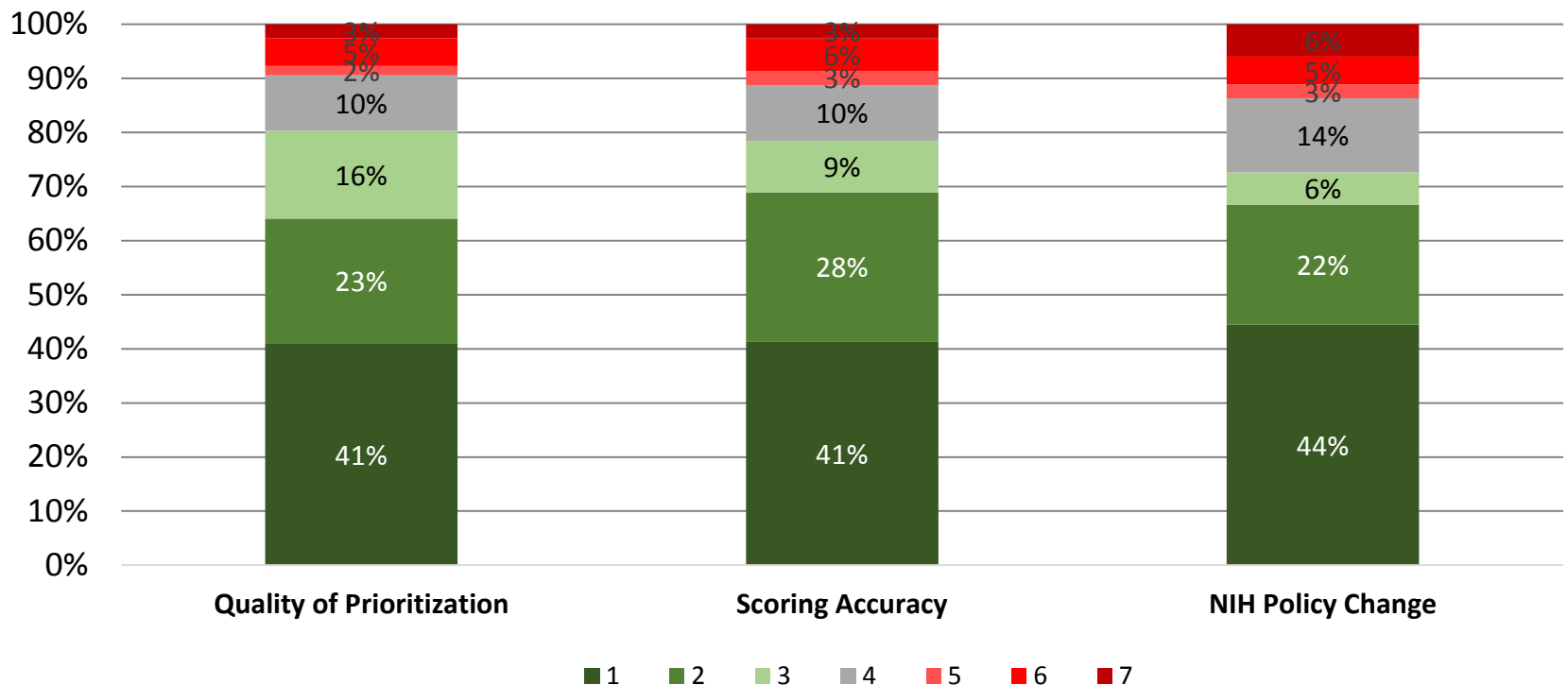
Conclusions

- Use of half-point option was used by over 80% of reviewers

- Effect of half-point option was to spread scores and reduce ties

- Two thirds of reviewers recommended a policy change

U.S. Department of Health & Human Services

Evaluation of NIH Peer Review Processes: The Effect of Anonymization

Lee S. Mann, Ph.D., J.D.

Scientific Review Officer, CSR

Anonymization Experiments – Basic Assumptions

- Large racial disparities in grant funding exists (Ginther et al): AAs have much lower award rates than Whites.

- At least 3 reviewers evaluate an application and the average preliminary overall impact scores are used to determine discussion order. 50% not discussed. Disparity of awards are determined by the final overall impact score.

- The major hypotheses for score disparity are:
    – Reviewer bias and/or
    – Quality of application submission

    ACD Rx's: CSR conduct studies using anonymization as a quality control check of our peer review process.

Specific Aims

1. **To determine if masking PII information from grant applications reduces the differences in final scores for Black and White applicants.**

2. To determine if this reduces the differences in final scores for Male and Female applicants. (secondary)

3. To determine if this reduces the differences in final scores for Established and ESI applicants. (secondary)

4. To determine if this reduces the differences in final scores for applicants from more research intensive and less research intensive institutions. (secondary)

Anonymizing Experiments – Design

1200 previously reviewed applications in 2014 – 2015 [400 AA, 400 Whites matched by science area, score, gender, degree, institution and seniority] & 400 Whites randomly selected

Application Formats		
A. Original R01 Application	B. Full Anonymization	
With Investigator and Institution Information	No applicant or institution information provided using entire application (Information will be redacted)	Test for differences in scores between Original and Full Anonymization

U.S. Department of Health & Human Services

Improving speed and accuracy of grant review and award

Improved Approaches to Review and Award?

- Retrospective review (R35 type) based on 3-5 publications
- No-deadlines for submissions
- Applicants are also reviewers of applications
- Editorial boards adjust order of priority from broad perspective
- Award decisions made within 6 months of submission

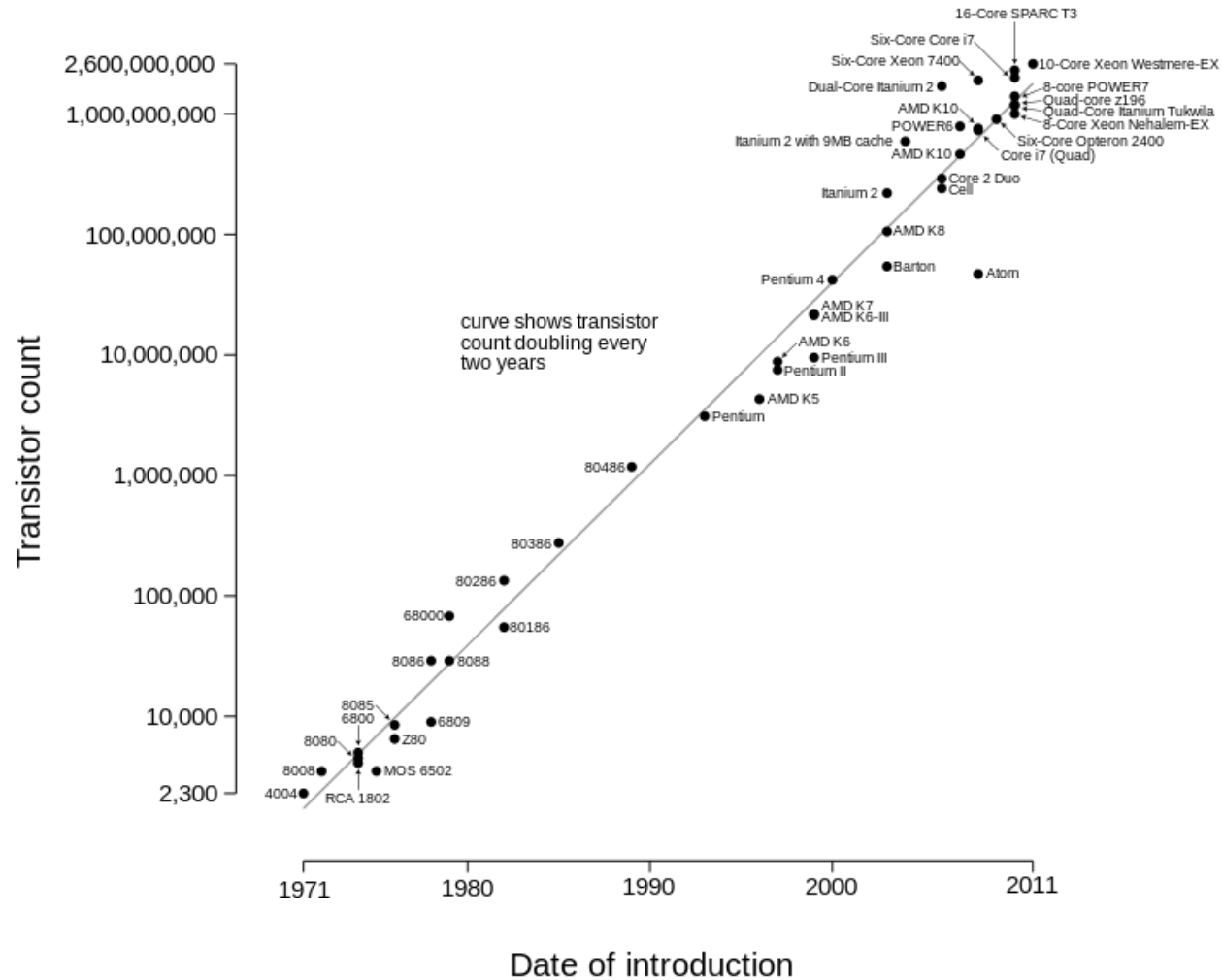- All modification to review tested before application

NIH Center for Scientific Review

Do We Have a Science and Technology Future?

Microprocessor Transistor Counts 1971-2011 & Moore's Law
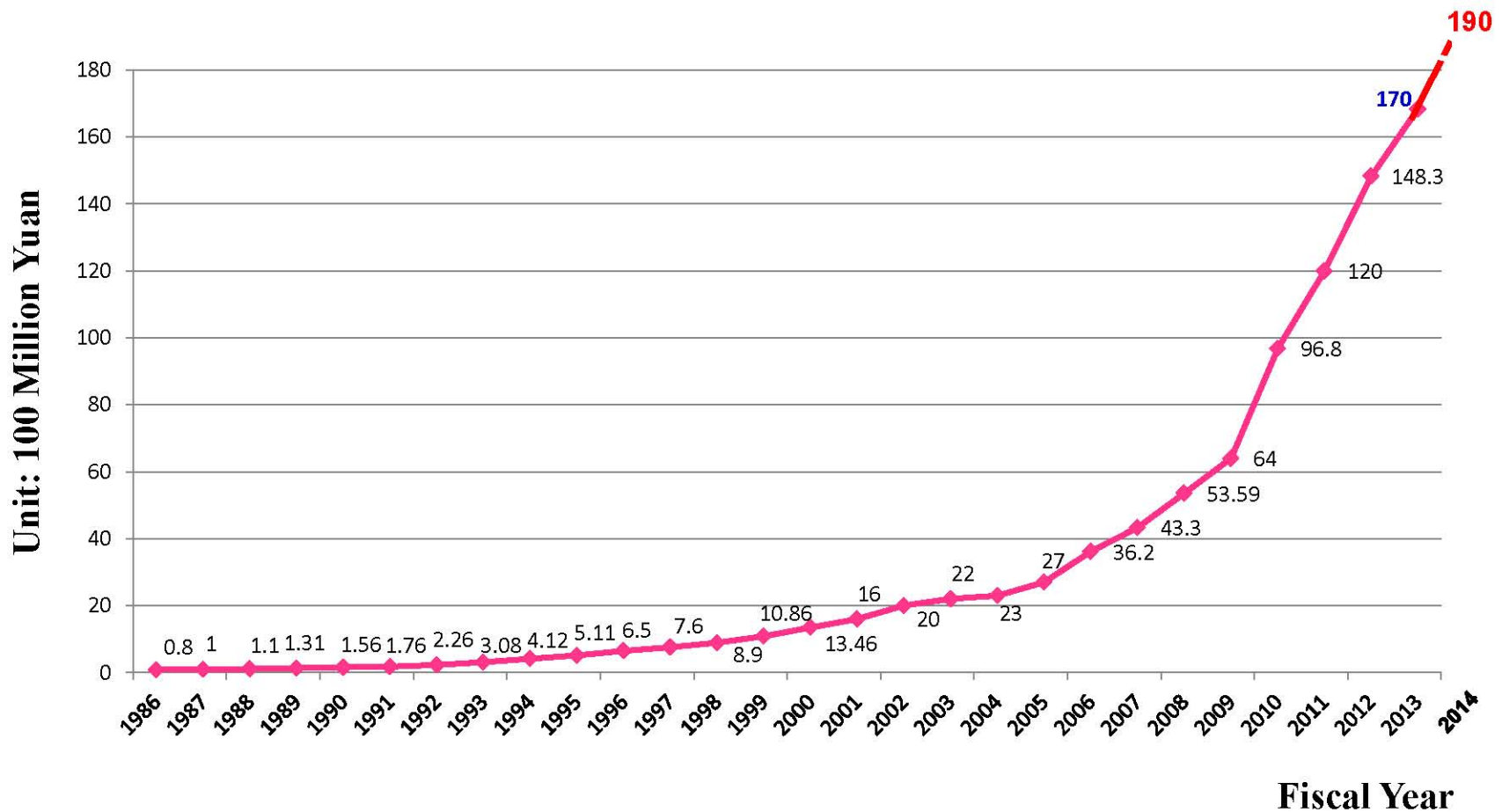
From Wikipedia attributed to user Wigsimon

What do these countries have in common?

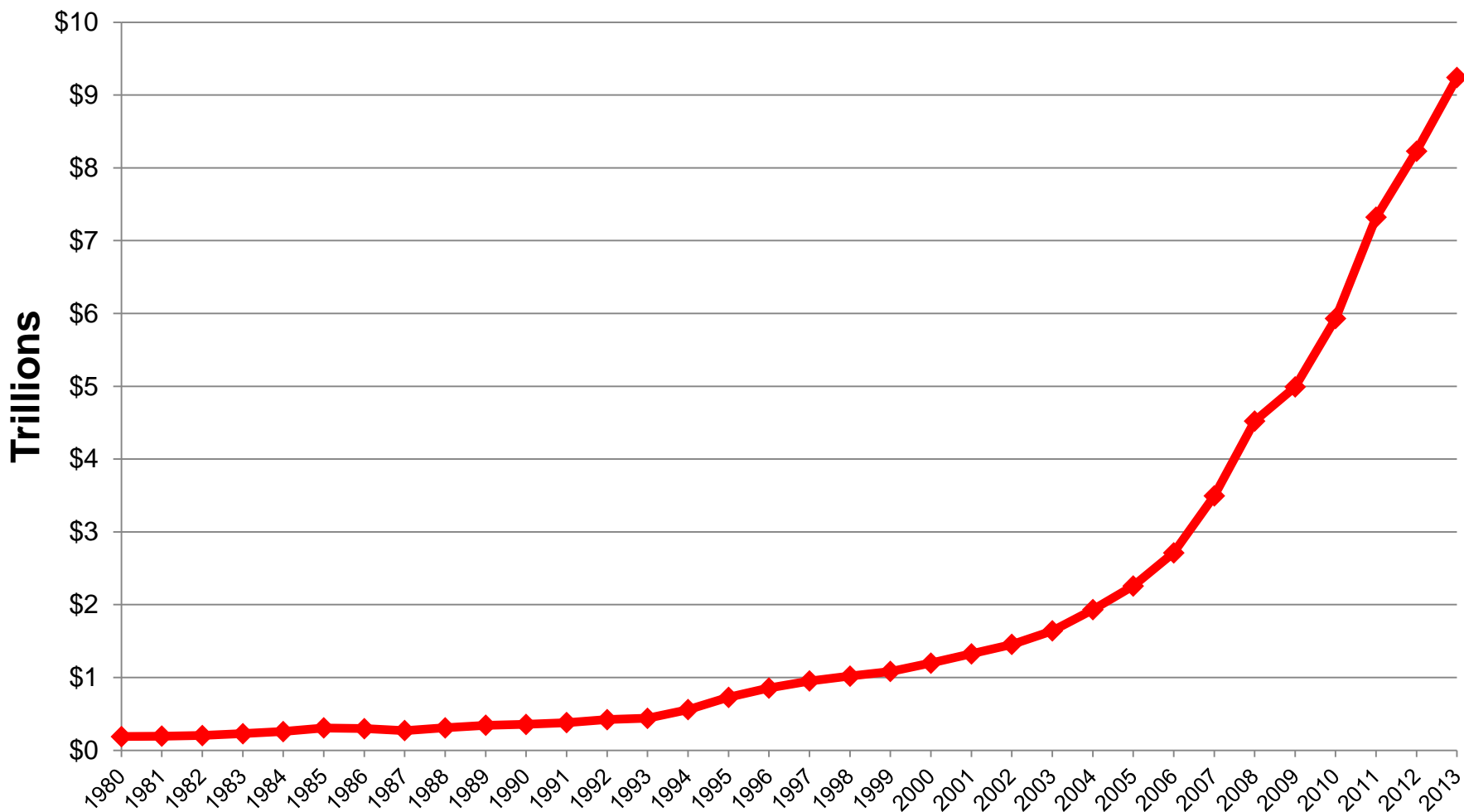- Israel   4.2
- South Korea   3.6
- Finland   3.5
- Japan   3.4
- Sweden   3.4
- Germany   2.9
- Switzerland   2.9
- Denmark   2.9

Budget for 1986-2014
The total budget for 2014 is ￥19Billion (~$ 3.05Billion), an increase by 11.7% over the year 2013.

China GDP in Current US Dollars, 1980-2013

http://data.worldbank.org/country/china

NIH Center for Scientific Review

Some happier thoughts

END