

Program Evaluation Criteria

Following enactment of the Government Performance and Results Act (GPRA) in 1993, the National Academies' Committee on Science, Engineering, and Public Policy produced a report on the unique purpose of federal research programs and inherent challenges in their evaluation. The committee concluded that federal research programs could be evaluated using three criteria: quality, relevance, and leadership, and noted that such evaluations should consider factors beyond peer review of research publications by scholars in the field (National Academy of Sciences, 2001).

In its 2008 Guide to the Program Assessment Rating Tool (PART) and citing the National Academies report, the US Office of Management and Budget (OMB) identified relevance, performance, and quality as criteria that can be used to assess the effectiveness of federal research and development (R&D) programs. This approach was further endorsed in a 2008 National Research Council (NRC) report, which stated that research program efficiency must be evaluated in the context of relevance, effectiveness, and quality.

NOAA, through an Administrative Order (NAO 216-115A, dated October 3, 2016, and its previous editions), has adopted Quality, Relevance and Performance as core evaluation criteria. The NAO also calls for a periodic evaluation of research, development and transition activities as well as outreach efforts and stakeholder engagement.

In general terms, these criteria may be described in the following terms:

- **Quality:** Quality is a measure of soundness, accuracy, and reproducibility of a specific body of research, and is the most widely and traditionally used criterion that is evaluated by peer review committees. In general, it refers to the merits of R&D within the scientific community – research publications, awards, innovations, and patents – and implies adherence to values of objectivity, fairness, and accountability (NAS, 2017). It also requires evidence of established procedures for competitive, merit-based procedures for research funding, and assuring scientific integrity.
- **Relevance.** In essence this criterion asks for “what would not have happened if NCCOS did not exist, and how much would society have missed?” Program personnel identify public benefits of the program, including added benefits beyond those of any similar effort that has been or is being funded by others. In this context, an “other” may be another program at the same or another agency, an interagency initiative or partnership, or a firm or other organization. R&D benefits include increasingly more skillful and reliable program output, technology, or methodology that satisfies legal mandates and user needs, expert counsel and technology transfer, and new options in the future.
- **Performance:** Performance refers to an ability to manage in a manner that produces identifiable results, both effectively (achieving desired results) and efficiently (with maximum productivity and minimum wasted effort or money). This is judged by program management structure that produces the desired results, guidance or framework for tracking progress toward agency’s strategic goals and objectives, flexibility to address events or changing priorities, interaction with stakeholders, and extramural collaboration.