Implementation Plan for Gold Standard Science: First Annual Report

July 23, 2025

Office of Policy Development & Research

Table of Contents

Introduction 1

Current Implementation of Gold Standard Science 2

Reproducible 2

Transparent 3

Communicative of Error and Uncertainty 4

Collaborative and Interdisciplinary 5

Skeptical of its Findings and Assumptions 5

Structured for Falsifiability of Hypotheses 6

Subject to Unbiased Peer Review 7

Accepting of Negative Results as Positive Outcomes 7

Without Conflicts of Interest 7

Plans for Further Implementation of Gold Standard Science 9

Revise HUD's Evaluation Policy Statement 9

Train HUD Leadership and Staff on Gold Standard Science 10

Further Integrate Gold Standard Science into Funding Opportunities 11

Leverage Technology to Enhance Gold Standard Science 11

Metrics and Evaluation Mechanisms to Assess Adherence to Gold Standard Science 13

Review of NOFOs and Contract Solicitations 13

Comprehensive Assessment of Completed Research Projects 13

Conclusion 16

Introduction

This report responds to Executive Order (EO) 14303, Restoring Gold Standard Science, dated May 25, 2025. EO 14303 builds upon the administration's broader agenda to reestablish objectivity, transparency, and accountability across federal policymaking. Aligned with EO 14151, which ended diversity, equity, and inclusion (DEI) mandates in government, and EO 14179, which promotes American leadership in artificial intelligence (AI) and emerging technologies, EO 14303 reinforces a return to evidence-based decisionmaking free from ideological influence. Together with the Executive Order on Preventing Woke AI in the Federal Government, these actions reflect a unified strategy to center scientific integrity, technological leadership, and neutral governance in federal operations.

EO 14303 defines scientific information as "factual inputs, data, models, analyses, technical information, or scientific assessments related to such disciplines as the behavioral and social sciences, public health and medical sciences, life and earth sciences, engineering, physical sciences, or probability and statistics. This includes any communication or representation of knowledge such as facts or data, in any medium or form, including textual, numerical, graphic, cartographic, narrative, or audiovisual forms."

Further, the EO defines *Gold Standard Science* as "science conducted in a manner that is: (i) reproducible; (ii) transparent; (iii) communicative of error and uncertainty; (iv) collaborative and interdisciplinary; (v) skeptical of its findings and assumptions; (vi) structured for falsifiability of hypotheses; (vii) subject to unbiased peer review; (viii) accepting of negative results as positive outcomes; and (ix) without conflicts of interest."

EO 14303 required the Director of the Office of Science and Technology Policy (OSTP) to issue guidance for agencies on the implementation of Gold Standard Science. OSTP issued this guidance in the form of a memorandum to the heads of executive departments and agencies on June 25, 2025. The guidance directs departments and agencies to implement Gold Standard Science "in the conduct and management of all aspects of their scientific activities, from research design to public communication."

The U.S. Department of Housing and Urban Development (HUD) welcomes EO 14303 and OSTP's guidance as an opportunity to build on the culture of rigor, transparency, and objectivity that characterizes the Department's production and use of scientific information. Consistent with OSTP's guidance, this report outlines HUD's implementation plans for Gold Standard Science, including—

- HUD's current implementation of Gold Standard Science.
- HUD's plans to further implement Gold Standard Science, including discussion of training, technology, and implementation challenges.
- Metrics and evaluation methods HUD can use to evaluate the implementation of Gold Standard Science.

Current Implementation of Gold Standard Science

For more than 50 years, the U.S. Department of Housing and Urban Development (HUD) has produced high-quality evidence to inform policy. Evidence-building occurs in several offices at HUD, but primary responsibility for producing scientific information lies with HUD's Office of Policy Development and Research (PD&R), whose mission is to improve life in American communities by building evidence to inform and strengthen housing and community development policy and practice and to accelerate progress on HUD's priorities. PD&R is responsible for most of the Department's research studies and program evaluations. PD&R also conducts economic and regulatory analysis for the Department on all new rules and regulations as required by Executive Order (EO) 12866 and approved by the Office of Management and Budget (OMB). Congress typically appropriates $10 million to $20 million each year specifically for research, evaluations, and demonstrations, which PD&R primarily conducts through cooperative agreements and contracts with external researchers. Additional appropriations fund the production of scientific data through surveys such as the American Housing Survey and the work of PD&R's staff of approximately 140 economists, social science analysts, and program analysts, who produce scientific information for the rest of HUD, other government agencies, researchers and policymakers, and the public.

Since its creation in 1973, PD&R has cultivated and sustained a national reputation for rigor and independence in its scientific activities and for effectively translating research into policy.[1] Today, PD&R's commitments to high scientific standards are documented in HUD's Evaluation Policy Statement, first published in 2016 and revised in 2021.[2] The Evaluation Policy Statement articulates six principles for HUD's evaluation work—rigor, relevance, transparency, independence, ethics, and technical innovation—and describes how HUD operationalizes those principles to produce and disseminate high-quality and consistent evaluation results. These six principles align with the nine tenets of Gold Standard Science as described in OSTP's guidance, though HUD's current Evaluation Policy Statement captures some of the tenets more fully than others. This section describes how HUD is currently addressing each of the tenets, including how the tenets are reflected in HUD's Evaluation Policy Statement and research culture, HUD's funding opportunities, award selection and reporting, and other actions relevant to the conduct and management of scientific activities. The discussion of each tenet begins with the definition, in italics, excerpted from OSTP's guidance.

Reproducible

> *Reproducibility in science is the ability of independent researchers to test a hypothesis through multiple methods and consistently achieve results that confirm or refute it, ensuring findings are generalizable and robust across different approaches. Replicability is the ability to perform the same experiment or study using the same methods and conditions to achieve the same result.*

Reproducibility and replicability of scientific results in the social sciences depend upon clear concepts and careful documentation of research methods. PD&R's research culture, reflected in HUD's Evaluation Policy Statement, places high value on clarity of conceptualization and careful documentation of research

[1] PD&R was created when the Department combined its previously separate research and policy functions to make research more relevant to policy and policy more grounded in empirical evidence.

[2] PD&R published HUD's first Evaluation Policy Statement in 2016 in response to a GAO assessment of HUD's management and program oversight that found the Department lacked agency-wide, written policies for its program evaluations or documented criteria for selecting expert panels and reviewing the quality of program evaluations. HUD updated the statement again in 2021 to align with the Foundations for Evidence-Based Policymaking Act of 2018 and OMB guidance on program evaluation standards.

methods. For example, the Evaluation Policy Statement states that "HUD is committed to using the most rigorous methods that are appropriate to the evaluation questions and feasible within budget and other constraints," with rigor characterized by "transparent, verifiable methods of systematic data collection and analysis [and] auditable records." Consistent with this commitment to verifiability and auditability, HUD's research studies, where feasible, include deliverables such as raw and final datasets and code for data cleaning and analysis that facilitate reproducibility and replicability.

Some of HUD's most important studies are too large and complex to be replicated in their entirety, but publishing resulting datasets allows the findings to be replicated and verified. For example, HUD's Moving to Opportunity demonstration randomized 4,600 participants to one of three housing conditions and lasted more than a decade—an experiment that would be impossible to replicate fully. However, a public use file with the study data was made available through the Inter-University Consortium for Political and Social Research (ICPSR), which permitted replication of the study's results. In recent years, HUD has also sponsored several replication studies, such as the Jobs Plus Replication Study (2021), which evaluated HUD's attempt to replicate the outcomes of the first round of Jobs Plus demonstration sites, and the Analysis of Housing Finance Issues Using the American Housing Survey (2003), which tested the reliability of the American Housing Survey by replicating measures of mortgage activity from other reliable sources of data.

Many studies that cannot be replicated in their entirety—such as the Moving to Opportunity Demonstration or the Effects of Housing Vouchers on Welfare Families study—lay the groundwork for future studies that build the evidence base by testing similar hypotheses with different approaches. Indeed, reproducibility has been a hallmark of HUD's research over the past few decades. HUD's research agenda supports multiple studies of the impact of housing assistance and services on key household outcomes, such as housing stability, health, employment, and earnings, and Congress historically has funded studies that examine the same types of outcomes in different contexts.[3] For example, since the 1990s, HUD has sponsored a series of studies that examine how the receipt of tenant-based rental assistance affects where households live and how much they work.

Transparent

> *Transparency in science entails the open, accessible, and comprehensive sharing of all components of the research process—methodologies, data, analytical tools, and findings—to enable stringent scrutiny, validation, and reuse by the scientific community and the public.*

HUD primarily achieves transparency by publishing its research and data products on the HUDUser.gov website. PD&R created HUD User in 1978 to distribute relevant research and information. In 2025, HUDUser.gov is a multimedia resource for housing stakeholders, with close to 45 million monthly hits. PD&R's research culture places a very high value on transparency and publication of HUD reports and data products. For example, HUD's Evaluation Policy Statement commits the Department to release methodologically valid evaluations without regard to the findings and to publish interim as well as final reports, explaining deviations from the original plans as applicable. The Evaluation Policy Statement also commits HUD to publishing synopses of the research design, data collection, and analysis plans for any studies that include a counterfactual and have broad public interest. PD&R's research publications typically include detailed methods sections and methodological appendixes, including disclosure of experimental protocols (if such exist) and analysis methods.

[3] HUD's research budget and priorities are determined by its research agenda and congressional appropriations and direction.

HUD's Evaluation Policy Statement states that HUD will archive administrative and evaluation data for use by interested researchers when possible. PD&R has a policy in place for allowing external researchers access to data through a data license or an interagency agreement. In cases in which research data collected are especially high quality (such as data produced through a randomized controlled trial or representative sampling methods) and useful to the general research community, HUD endeavors to make the data available in a public use file through vehicles such as the U.S. Census Bureau's Data Linkage Infrastructure and ICPSR at the University of Michigan. Funding for research contracts and cooperative agreements may include support for the grantee or contractor to prepare and document data files for public use.

To support the implementation of EO 14303, HUD's Office of the Chief Information Officer (OCIO) is coordinating efforts to ensure that technology systems align with the principles of Gold Standard Science. This work includes facilitating the public disclosure of data, models, and source code; improving data traceability; and promoting the use of advanced tools, including AI, to enhance reproducibility and transparency. OCIO is also leading the development of internal mechanisms to monitor scientific integrity compliance, consistent with federal open data and evidence-building policies.

Transparency also extends to HUD's processes for funding research through grants and contracts. The evaluation criteria used to review and score applications for cooperative agreements and contract bids are published as part of the Notices of Funding Opportunity (NOFOs) (cooperative agreements) or the Request for Contract Services (contracts). The evaluation criteria for grants and contract awards include requirements for data sharing if appropriate and feasible. HUD also has procedures for debriefing unsuccessful applicants for cooperative agreements and contract bidders upon request.

Communicative of Error and Uncertainty

> *Communicating error and uncertainty in science entails the clear, precise, and accurate disclosure of limitations, variability, and potential sources of error or limitations in measurements or research findings, enabling other scientists to critically assess, replicate, and extend the work.*

PD&R's research culture places a high value on clearly communicating error and uncertainty in relation to all studies, including program evaluations, policy-relevant research, and economic analysis of regulations, and strongly discourages speculative claims or extrapolations that extend beyond the data's scope. This culture is reflected in the Evaluation Policy Statement, which states that "evaluation reports must describe the methods used, including strengths and weaknesses, and discuss the generalizability of the findings" and, for economic analyses of regulations, that HUD uses and produces "the best objective estimates of the benefits, costs, and transfers resulting from the regulation, taking into account gaps and uncertainties in the available data and methodologies." When they review research designs and reports, PD&R staff look for information—such as confidence intervals and measures of multicollinearity—that can affect the validity of the analysis and interpretation of results. PD&R staff work with researchers to ensure that this information is clearly presented in the report or its technical appendixes.

HUD communicates error and uncertainty in its reports and data products by providing clear explanations of methodological constraints and assumptions and reporting confidence intervals, error margins, p-values, and results of sensitivity analyses. For example, the Long-Term Impact Report from HUD's First-Time Homebuyer Education and Counseling Demonstration includes multiple appendixes that detail the study's analytic methods, reports on sensitivity analyses the study team conducted using alternative outcome specifications, presents the findings for *all* impact analyses conducted for the study (not just those featured in the main body of the report), and offers clear instructions for the lay reader on how to interpret impact

estimates, confidence intervals, minimum detectable effects, and other key concepts for interpreting the study results.

Collaborative and Interdisciplinary

> *Collaborative and interdisciplinary science refers to the strategic integration of a wide range of expertise, methodologies, and perspectives across disciplines and sectors to address complex scientific challenges and catalyze transformative discoveries.*

Collaboration and interdisciplinarity are fundamental components of the research culture at HUD, through PD&R. PD&R's mission is to build "evidence to inform and strengthen housing and community development policy and practice and accelerate progress on HUD's priorities." This mission requires drawing on a range of disciplines, from architecture to urban studies. PD&R research staff come from a variety of disciplines spanning the humanities; economics and finance; architecture and planning; the social, biological, and natural sciences; and engineering. PD&R research staff collaborate with HUD program offices and external entities to gather input into HUD's Learning Agenda and project-specific research designs. Funding opportunities are not restricted to single disciplines and often require multidisciplinary teams. Grants and contracts for research often include an external panel of experts from several disciplines to comment on key project deliverables, such as the research design, data collection and analysis plan, and substantive reports.

In addition to seeking opportunities to bring together interdisciplinary teams, PD&R staff collaborate with colleagues across offices and federal agencies to benefit from their expertise or data resources. For example, PD&R works regularly with the Office of Lead Hazard Control and Healthy Homes at HUD and has collaborated with the U.S. Department of Health and Human Services to inform a research agenda on housing and health. PD&R enters into data-sharing agreements to support studies on the health of HUD-assisted households, such as A Health Picture of HUD-Assisted Adults, 2006–2012, which resulted from collaboration between HUD and the National Center for Health Statistics. In another example of collaboration, PD&R deposits data from its large randomized controlled trials with the Census Bureau's Data Linkage Infrastructure, where the data are available for linkage and analysis.

HUD supports internal and external collaboration by ensuring that HUD staff have access to collaborative digital tools, including platforms that can be used to facilitate real-time communication, data sharing, and document sharing. Ensuring that HUD staff and research stakeholders have access to centralized data storage solutions also ensures the timely transfer of knowledge.

Skeptical of Its Findings and Assumptions

> *Maintaining constructive skepticism of findings and assumptions in science refers to the critical and open-minded evaluation of research findings, methodologies, and underlying assumptions to ensure their validity, robustness, and reliability.*

HUD's research staff routinely engage in constructively critical and deliberately skeptical discussions of research, including staff's own work. PD&R's research culture emphasizes the need for constructive criticism and a skeptical eye throughout the research process. The teams that rate and review applications for cooperative agreements and offeror bids for contracts typically engage in healthy debate and open-minded evaluation. However, the importance of maintaining constructive skepticism of findings and assumptions is not explicit in HUD's Evaluation Policy Statement and should be added to codify and strengthen this aspect of PD&R's research approach.

Cityscape, PD&R's multidisciplinary journal of housing and community development, plays a role in fostering debate and constructive skepticism. As PD&R's Assistant Secretary Michael Stegman wrote in his introductory essay to the journal's third volume, "*Cityscape* began publication in 1994 with a commitment to present research that would 'challenge long-held assumptions . . . and encourage innovative thinking' on the crucial housing and urban development issues facing our Nation." To that end, *Cityscape* regularly features symposia presenting different perspectives on a topic, for example, "Federalism and Flexibility: Fifty Years of Community Development Block Grants." *Cityscape* also sometimes publishes a "Point of Contention," in which scholars with differing viewpoints are encouraged to respond to a common question, such as whether eviction can be accurately measured in the American Housing Survey. As part of its further implementation of Gold Standard Science, PD&R will assess the value and feasibility of creating a new department in *Cityscape* oriented around constructive skepticism.

Structured for Falsifiability of Hypotheses

> *Structuring science for falsifiability of hypotheses entails designing research studies and experiments to enable hypotheses to be carefully tested and potentially disproven through empirical evidence.*

When appropriate, given the state of knowledge and the research questions to be answered, HUD (through PD&R) promotes research and analysis methods to enable falsification of hypotheses. For example, when it is appropriate and feasible to use experimental methods, PD&R prioritizes doing so. Per the Evaluation Policy Statement, "In assessing the effects of programs or services, HUD evaluations use methods that isolate to the greatest extent possible the impacts of the programs or services from other influences, such as trends over time, geographic variation, or pre-existing differences between participants and non-participants. Where feasible, research should employ a treatment group and a counterfactual." Randomized controlled research designs create counterfactuals so that observers may ascertain whether outcomes would have been the same without the intervention, thereby enabling falsification of hypotheses.

HUD's Evaluation Policy Statement commits HUD to publishing synopses of the research design, data collection, and analysis plan for any studies that include a counterfactual and have broad public interest, and to publishing an explanation if the research deviates from the original plan. Publishing the research design synopses is a way to preregister study protocols. PD&R encourages the use of statistically valid sampling methods and quasi-experimental designs that allow for confirmation or rejection of hypotheses and support the identification of a small set of confirmatory outcomes at the onset of a study and a clearly specified larger set of exploratory outcomes to avoid the multiple comparison problem. In addition, HUD's Evaluation Policy Statement states that "[w]hen statistically appropriate, and particularly if the number of hypotheses being tested is large, HUD will require the use of commonly accepted adjustments to classical statistical testing to reduce the probability that random outliers are presented as meaningful."

Recent HUD studies with randomized controlled research designs or experimental or quasi-experimental analysis include Moving to Opportunity; Welfare to Work Voucher Program Evaluation; Family Options Study; Small Area Fair Market Rent Demonstration Evaluation; Supportive Services Demonstration; Housing and Children's Healthy Development Study; Rent Reform Demonstration; Choice Neighborhoods Evaluation; Family Self-Sufficiency Program Evaluation; HUD First-Time Homebuyer Education and Counseling Demonstration; Evaluation of the Community Choice Demonstration; and Evaluation of the Moving to Work Demonstration Expansion.

Subject to Unbiased Peer Review

> *Subjecting science to unbiased peer review (sometimes referred to as merit review) refers to the impartial and independent evaluation, by qualified experts, of both research proposals and manuscripts that report results of federally supported research, to ensure validity, quality, and credibility prior to funding, publication, or dissemination.*

HUD uses unbiased peer review to review, select, and award federal grants and contracts. Research proposals undergo independent, impartial peer review guided by clear, transparent evaluation criteria. The review process is standardized to ensure objectivity and consistency. Reviewers are selected on the basis of expertise and independence. Reviewers for grants must be able to attest that they have no conflicts of interest, complete a financial disclosure report, and commit in writing to avoid disclosure of material used in the review process. Contracts incorporate the Federal Acquisition Regulation clause 52.203-16: Preventing Personal Conflicts of Interest, which covers reviewers of proposals. Reviewers are instructed to provide unbiased technical reviews in relation to the published rating factors and to recommend awards based on merit. Rating factors and evaluation criteria are published with each Request for Contract Services and each NOFO. For both contracts and cooperative agreements, reviews and recommendations are documented and stored. Many of the articles in HUD's journal *Cityscape* also undergo peer review before acceptance and publication.

Accepting of Negative Results as Positive Outcomes

> *Accepting negative results as positive outcomes in science refers to recognizing and valuing—as meaningful contributions to knowledge generation—null or unexpected findings that fail to support a hypothesis.*

HUD's research culture values the publication of all methodologically valid results, including negative results. HUD's Evaluation Policy Statement states that "HUD will release methodologically valid evaluations without regard to the findings" and that "[e]valuation reports must present comprehensive results, including favorable, unfavorable, and null findings." HUD has a track record of publishing reports with null findings for some or all the outcomes tested. Recent publications with null findings on key hypotheses include the Rent Reform Demonstration, the Family Self-Sufficiency Program Evaluation, and the HUD First-Time Homebuyer Education and Counseling Demonstration.

Without Conflicts of Interest

> *Conducting science without conflicts of interest refers to ensuring that research is designed, executed, reviewed, and reported free from financial, personal, or institutional influences that could bias outcomes or undermine objectivity.*

HUD's research culture acknowledges and prohibits conflicts of interest. HUD staff involved in defining and developing the Department's research agenda are subject to HUD's conflict of interest regulations. PD&R staff collaborate with external partners to collect, summarize, and review identified research needs and priorities from a wide range of stakeholders, but research funding is allocated through peer-review processes (as described in previous sections of this report).

HUD allocates research funding partly through contracts and cooperative agreements. When allocating research funding, reviewers comply with all laws and regulations applicable to the procurement of research through contracts and to federal grantmaking. For example, before any contract review process, contracting officers (COs) ensure that applicable participating staff (for example, members of Technical

Expert Panels and Source Selection Authority) certify that they have no conflicts of interest via formal paperwork (e.g., form SF 278 for presidential appointees and senior executives and form OGE 450 for career staff). When conflicts are identified, the Agency Ethics Officer reviews all reported financial conflicts of interest. COs are responsible for ensuring that the Agency Ethics Officer has reviewed and certified all disclosure forms before finalizing the review process. During the review process, several mechanisms are employed to avoid conflict of interest issues. For example, proposal information is not shared with Technical Expert Panel advisors. If an advisor identifies a conflict upon learning of proposing firms, they must immediately notify relevant parties, stop advising the panel, and comply with any disclosure reporting duties and requirements.

To ensure that HUD's staff comply with conflict of interest regulations and processes, staff maintain required certifications, such as the Contracting Officer's Representative (COR) certification, and participate in regular trainings about ethics, procurement, and federal grantmaking. HUD's Evaluation Policy Statement commits HUD to conduct "evaluations through the competitive award of grants and contracts to external experts who are free from conflicts of interest."

Plans for Further Implementation of Gold Standard Science

The review of HUD's current implementation of Gold Standard Science suggests that HUD, primarily through the Office of PD&R, has a research culture, policies, and practices that align with Gold Standard Science. However, for each of the tenets of Gold Standard Science, there are places where HUD could be doing more. HUD plans to further the implementation of Gold Standard Science across the Department through four main actions:

1. Revise HUD's Evaluation Policy Statement.
2. Train HUD leadership and staff on Gold Standard Science.
3. Further integrate Gold Standard Science into funding opportunities as appropriate.
4. Leverage technology to enhance Gold Standard Science.

Each action is described below, including discussion of anticipated challenges and barriers.

Revise HUD's Evaluation Policy Statement

HUD plans to revise its Evaluation Policy Statement to more fully incorporate the nine tenets of Gold Standard Science and document existing HUD practices that support Gold Standard Science but are not explicit in the Evaluation Policy Statement. HUD will take a closer look at its current policies and practices related to producing and using scientific information and how they address the tenets of Gold Standard Science, building from the work conducted for Section 1 of this plan. PD&R will engage producers and consumers of scientific information across HUD, as well as those managing the Department's technology, for this analysis. The analysis is expected to yield ideas for new practices that can be documented in the revised Evaluation Policy Statement.

Exhibit 1 outlines potential modifications to the Evaluation Policy Statement for each of the tenets based on the review conducted to date. In areas where HUD and PD&R's existing policies and practices are especially strong, the current recommendation is simply to review the existing language in the Evaluation Policy Statement for consistency with Gold Standard Science and make modest revisions if needed.

The main challenge to revising the Evaluation Policy Statement will be prioritizing staff time to review existing policies and practices and develop the new policy language.

Exhibit 1. Potential Modifications to HUD's Evaluation Policy Statement by Tenet

Tenet	Potential Modifications
Reproducible	Revise the existing language to be more explicit about the importance of facilitating and incentivizing reproducibility and replicability in projects that test a hypothesis (i.e., not descriptive or exploratory research).
Transparent	Review existing language for consistency with Gold Standard Science and revise if needed, consistent with applicable confidentiality and privacy standards and laws. Cite standardized metadata formats as applicable, as well as requirements for data sharing or linkage for certain types of research.
Communicative of error and uncertainty	Review existing language for consistency with Gold Standard Science and revise if needed.
Collaborative and interdisciplinary	Add more language to reflect HUD's existing practices around collaborative and interdisciplinary work and to increase attention on team science.

Tenet	Potential Modifications
Skeptical of its findings and assumptions	Add more language about the importance of a culture of constructive skepticism of findings and assumptions.
Structured for falsifiability of hypotheses	Review existing language for consistency with Gold Standard Science and revise if needed. Consider a more formal approach to preregistering study protocols for randomized controlled trials.
Subject to unbiased peer review	Add more language to clarify existing peer-review procedures and assess whether additional types of activities should be subject to unbiased peer review.
Accepting of negative results as positive outcomes	Revise the departmental clearance process for research findings to withhold only results that are inconsistent with Gold Standard Science and administration priorities.
Without conflicts of interest	Review existing language for consistency with Gold Standard Science and revise if needed.

Train HUD Leadership and Staff on Gold Standard Science

HUD plans to provide a series of training opportunities on Gold Standard Science following the development of the revised Evaluation Policy Statement. HUD envisions developing training materials for both *consumers* of scientific information within HUD, including offices responsible for public communication of such information, and *producers* of scientific information within HUD.

Introductory Training

HUD plans to produce an introductory training course on Gold Standard Science with the goal of ensuring that HUD staff and leadership are aware of EO 14303, OSTP's guidance, and HUD's responsibilities for complying with the EO and guidance. The training would be offered at multiple times and recorded to ensure that all staff have access to it. HUD will consider options to guarantee that all HUD personnel take the training, such as making it part of the mandatory annual training.

The introductory training will cover the following concepts:

- The application of EO 14303 to all aspects of HUD's scientific activities.
- The nine tenets of Gold Standard Science.
- HUD's revised Evaluation Policy Statement.
- HUD's plans for implementation of Gold Standard Science.
- The requirement for annual reporting on HUD's implementation of Gold Standard Science.
- EO 14303's directive regarding the use, interpretation, and communication of scientific data.

The introductory training would be the same for all HUD personnel—both consumers and producers of scientific information.

In-depth Training for Producers of Scientific Information

In addition to the universal training, HUD personnel directly involved in producing scientific information will receive in-depth training on the core tenets of Gold Standard Science. This training would apply to anyone producing scientific information, either directly producing the information or managing extramural contracts and cooperative agreements. The training would be structured in modules and would cover each of the nine tenets thoroughly using HUD-specific examples. Following are some examples of what the in-depth training could cover for a given tenet:

- **Reproducible.** How HUD staff producing and managing research can look for ways to ensure that studies can be replicated and/or reproduced to the extent feasible within budget and other constraints.
- **Transparent.** What options HUD staff have for making data and code from scientific activities accessible to the public, with appropriate privacy protections and oversight to review the materials for completeness and accessibility.
- **Communicative of Error and Uncertainty.** Examples of standardized formats for reporting uncertainty in social science research and strategies for ensuring that uncertainty is plainly communicated (and not minimized) when HUD uses scientific information in ways that influence program operations or policy, such as in proposed regulations or regulatory impact analyses.
- **Subject to Unbiased Peer Review.** The norms and practices of peer review and which aspects of HUD's production of scientific information either require or would benefit from peer review.

The main challenge to providing the above training opportunities on Gold Standard Science will be identifying the staff and other resources to assemble and deliver the training, as well as developing approaches to help ensure that the relevant staff and leadership make time to take the training.

Further Integrate Gold Standard Science into Funding Opportunities

As part of the review preceding the revision of the Evaluation Policy Statement, HUD will consider ways to integrate Gold Standard Science further into its funding opportunities related to the production of scientific information.

One idea is to require that program descriptions in NOFOs for cooperative agreements and statements of work in solicitations for contracts reference EO 14303 and the tenets of Gold Standard Science, as well as HUD's Evaluation Policy Statements. For research projects intended to test a hypothesis, the program descriptions or statements of work could potentially include requirements related to delivering data and codes to facilitate replicability or reproducibility and requirements for standardized communication of error and uncertainty.

HUD could also review its policies and procedures for monitoring potential conflicts of interest of reviewers, researchers, and agency officials involved in the funding or performance of federal research to identify areas for improvement. HUD could further assess whether requirements for disclosure of financial, personal, or institutional interests in research proposals, publications, peer and merit reviews, and data repositories should be enhanced.

As with the other components of HUD's plan to further its implementation of Gold Standard Science, the main challenge to reviewing its funding opportunities will be prioritizing staff resources to do the work.

Leverage Technology to Enhance Gold Standard Science

HUD is positioning itself well to meet the technology expectations of EO 14303 and related directives through the leadership of its Chief Technology Officer (CTO) within the Office of the Chief Information Officer (OCIO). Through the cooperation of the Chief Information Officer and the Chief Data Officer (CDO), HUD ensures that modern, scalable technologies directly support the production and dissemination of scientific information. This structure enables close collaboration across all program

offices—including PD&R—to promote transparency, reproducibility, and public access to data, models, and evidence-based tools.

Implementing HUD's Fiscal Year 2026 Performance Plan

HUD recognizes that outdated processes must be replaced with upgraded technology solutions. In HUD's Fiscal Year 2026 Annual Performance Plan (APP), the Department outlines plans to leverage technology to "radically streamline and simplify HUD operations to meet 21st Century demands." APP Priority 4—Streamline Business Processes—discusses a strategic restructuring of IT processes at HUD to optimize efficiency and effectiveness. To ensure success, HUD established a key performance indicator focused on leveraging technology to streamline processes using IT modernization.[4] HUD intends to track "the number of HUD processes that have been automated, improved, or replaced through the implementation of IT modernization techniques." Through streamlining access and the use of better technology, HUD stakeholders can better support Gold Standard Science. HUD also publishes an annual IT expenditure plan that outlines specific technological investments. HUD's FY2026 Office of the Chief Information Officer budget request includes $365 million for the Information Technology Fund, which includes Development, Modernization, and Enhancement activities. Appropriations language focuses on investing in cybersecurity, improving data management processes, and enhancing financial and grants management systems. All these modernization efforts support the use of technology to better implement Gold Standard Science. The two primary challenges associated with implementation are budgetary constraints and staff capacity.

Building on HUD's Open Data Policy

HUD's CDO, in coordination with the CTO and PD&R teams, is leading the implementation of HUD's Open Data Policy. This policy uses modern technology to advance transparency, reproducibility, and replicability across the agency's scientific activities. As part of this effort, HUD will ensure that metadata related to HUD's research priorities and funded projects are published and maintained. When appropriate, OCIO will also oversee the development of a public repository for datasets and/or code generated through funded research. All publicly released materials will undergo review to ensure that privacy protections, oversight compliance, and accessibility standards are met. One recent example of PD&R-funded research that can be used as a gold standard is the County Emergency Rental Assistance Spending project. The project team published a publicly available repository for the county-level aggregated data using a technology product commonly used in private industry. Future studies should consider whether project deliverables should include datasets and/or code used for analysis. The primary challenge associated with implementing HUD's Open Data Policy is the balance between ensuring that data are publicly available while also adhering to appropriate data privacy requirements.

[4] One example of IT modernization relevant to the implementation of Gold Standard Science could be the use of virtual data enclaves for more efficient and timely data sharing.

Metrics and Evaluation Mechanisms to Assess Adherence to Gold Standard Science

After revising the Evaluation Policy Statement and training agency staff and leadership on Gold Standard Science, HUD will assess adherence to the tenets of Gold Standard Science and their impact on scientific quality using standardized metrics and evaluation mechanisms. PD&R is considering two approaches: (1) annual review of program descriptions and statements of work in NOFOs and contract solicitations and (2) comprehensive assessment of completed research projects. The first approach measures the extent to which HUD has integrated the language and tenets of Gold Standard Science into its funding mechanisms to ensure that applicants for funding expect to be required to follow the Gold Standard Science tenets to the extent feasible and that HUD has a formal mechanism to ensure that recipients of research funds do so. The second approach documents the extent to which the research conducted and its publications meet the standards set forth in OSTP's guidance, which has direct implications for the quality of the resulting scientific information.

Review of NOFOs and Contract Solicitations

HUD would annually review all NOFOs and contract solicitations published in the previous year related to the production of scientific information for references to EO 14303 and the tenets of Gold Standard Science. The review could be conducted using artificial intelligence (AI). The simple metric could be the percentage of NOFOs and statements of work that include language requiring applicants to adhere (to the extent feasible) to the tenets of Gold Standard Science in designing and executing their research. The goal would be to reach 100-percent adherence once HUD has incorporated the Gold Standard Science language into its policies and practices.

Comprehensive Assessment of Completed Research Projects

HUD would annually review all completed research projects funded through a HUD research NOFO or contract solicitation. The review would use a structured review guide and assign points for adherence to each of the nine tenets of Gold Standard Science. This review could potentially use AI to assemble much of the information and produce a preliminary assessment, which would then be reviewed and confirmed by HUD's Evaluation Officer or delegate.[5]

Exhibit 2 provides a starting point for an assessment tool that could be used consistently across projects to produce an overall score for each project, a score by tenet for each project, and average overall scores and scores by tenets across all completed projects. The goal would be for the average overall score, as well as the score for each tenet, to increase over time until an average score is achieved that represents excellent implementation of Gold Standard Science.

[5] The reviewer should be someone at arms' length from the projects to ensure an impartial review.

Exhibit 2. Outline for Comprehensive Assessment of Completed Research Projects

Tenet	Areas for Evaluation	Scoring Considerations
Reproducible	• Did the study produce sufficient information (methods, data, code, etc.) for the study findings to be reproduced using other methods? • Did the study produce sufficient information (methods, data, code, etc.) for the study findings to be replicated using the same methods? • Is the information needed to reproduce or replicate the study publicly available?	Metrics would apply only to projects intended to test a hypothesis (i.e., not descriptive or exploratory research). Points awarded for each "yes" answer; 0 points for "no" answers, and explanation required.
Transparent	• Does the study have a published Research Design? • Does the study have a published Data Collection and Analysis Plan? • Were the study's interim findings (if applicable) published? • Were the study's final findings published?	Points awarded for each "yes" answer; 0 points for "no" answers, and explanation required.
Communicative of Error and Uncertainty	• Does the study's final report (or related materials)— o Report data ranges, not just measures of central tendency? o Quantify statistical uncertainties? o Document and report potential sources of error? o Clearly articulate assumptions and methodological limitations? o Disclose potential biases? • Do public-facing materials about the study (executive summary, press releases, etc.) use cautious, evidence-based language in presenting the findings?	Points awarded for each "yes" answer; 0 points for "no" answers, and explanation required.
Collaborative and Interdisciplinary	• To what extent did this research take a collaborative and interdisciplinary approach (e.g., drawing on staff in other offices within the agency, partnering or sharing data with another agency, organizing an interdisciplinary expert panel, setting up an interdisciplinary research team)? o Highly collaborative and/or interdisciplinary (score = 2) o Moderately collaborative and/or interdisciplinary (score = 1) o Minimally collaborative and/or interdisciplinary (score = 0)	Requires reviewer assessment of extent of collaboration and interdisciplinarity (vs. yes or no answer).

Tenet	Areas for Evaluation	Scoring Considerations
Skeptical of Its Findings and Assumptions	• Did the study feature validation methods such as peer and merit review, replication studies, sensitivity analyses, or uncertainty assessments? o Multiple methods (score = 2) o One method (score = 1) o No validation method (score = 0)	Requires quantification of validation methods.
Structured for Falsifiability of Hypotheses	• Did the study formulate precise, testable hypotheses? • Did the study employ methodologies that were rigorous and appropriate for the hypotheses? • Did the study feature controlled experiments, randomized trials, or advanced statistical tests? • Did the study preregister its protocols and outcome measures?	Metrics would apply only to projects intended to test a hypothesis (i.e., not descriptive or exploratory research). Points awarded for each "yes" answer; 0 points for "no" answers, and explanation required.
Subject to Unbiased Peer Review	• Was peer review used to award the grant or contract? • Were the review criteria transparent and well defined? • Did the reviewers disclose conflicts of interest? • Were the review results available to applicants?	Points awarded for each "yes" answer; 0 points for "no" answers, and explanation required.
Accepting of Negative Results as Positive Outcomes	• Were the study findings published? o Completely, including negative or null results (score = 2) o Partially; some negative/null results not published (score = 1) o Study findings were not published or no negative/null results were published (score = 0)	Metrics would apply only to projects intended to test a hypothesis (i.e., not descriptive or exploratory research). Requires identification and characterization of unpublished results.
Without Conflicts of Interest	• Did all individuals and entities in the funding or performance of the study receive training on conflicts of interest and have a venue for disclosing conflicts of interest? o Grant application/proposal reviewers o Managers of research o Peer reviewers of research products	Points awarded for each "yes" answer; 0 points for "no" answers, and explanation required.

Conclusion

HUD welcomes the publication of Executive Order 14303 and the Office of Science and Technology Policy's implementation guidance. The EO presents an opportunity for HUD to build on the culture of rigor, transparency, and objectivity that has characterized the Department's production and use of scientific information for more than 50 years. This first annual plan provides a preliminary review of HUD's current adherence to the nine tenets of Gold Standard Science and offers a plan for furthering implementation of the tenets by revising HUD's Evaluation Policy Statement, training HUD leadership and staff, further integrating Gold Standard Science into funding opportunities, and leveraging technology. The main challenge to fulfilling this plan will be identifying the staff and resources needed to complete a thoughtful examination across all nine tenets and draft revised policies and procedures. HUD's Office of PD&R and Office of CIO will need to take the lead in implementing HUD's plan. Notwithstanding these challenges, HUD looks forward to refining its plan, implementing some of the proposed actions, and measuring its progress toward implementing Gold Standard Science using a standardized and objective approach.