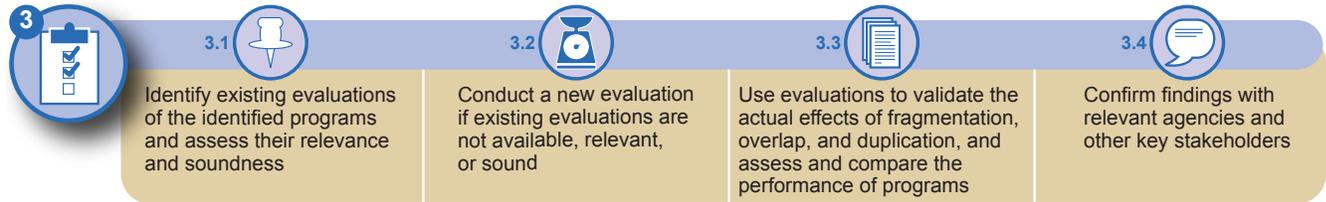


Step 3: Validate Effects and Assess and Compare Federal Programs

This step outlines for analysts how to use the results of existing or new evaluations to validate the effects of fragmentation, overlap, or duplication identified in Step 2 and assess and compare the programs identified in Step 1 to determine their relative performance and cost-effectiveness. Information gathered in this step can help analysts determine which programs or aspects of programs are more efficient and effective and identify options to reduce or better manage fragmentation, overlap, or duplication as outlined in Step 4 of this guide.



3.1 Identify existing evaluations of the identified programs and assess their relevance and soundness.

Analysts should first identify existing evaluations on the fragmented, overlapping, or duplicative programs identified in Step 1 of this guide. A program evaluation is a systematic study using research methods to collect and analyze data to assess how well a program is working and why. Evaluations answer specific questions about program performance and may focus on assessing program operations or results. Evaluation results may be used to assess a program's effectiveness, identify how to improve performance, or guide resource allocation.¹⁶

The following types of organizations may conduct program evaluations:

- a program office within the administering agency;
- a policy or evaluation office within the administering agency;
- an oversight-related agency, such as GAO or an agency's inspector general;
- a congressional or executive branch research or oversight agency, such as CRS or OMB;
- an independent consulting firm or research institute; or
- an academic institution.

To determine whether an organization has conducted evaluations of the identified programs, analysts can review existing sources of information, such as performance reports required by the Government Performance and Results Act of 1993 (GPRA), as updated by the GPRA Modernization Act of 2010 (GPRAMA), or conduct original research, such as interviews with agency officials.¹⁷

If analysts identify existing evaluations on the selected programs, they should assess the purpose, scope, and design of existing evaluations to determine if they are relevant to the areas where potential positive or negative effects were identified in Step 2 (**implementation, outcomes and impact, or cost-effectiveness**).

- **Purpose** establishes the questions an evaluation will and will not answer.
- **Scope** defines subject matter of the evaluation (such as part of a program, an entire program, or

¹⁶See GAO, *Designing Evaluations: 2012 Revision*, [GAO-12-208G](#) (Washington, D.C.: January 2012).

¹⁷Pub. L. No. 111-352, 124 Stat. 3866 (2011) (GPRAMA). GPRAMA requires agencies to summarize in their annual performance reports any program evaluations that were conducted during the past year. 31 U.S.C. § 1116(c)(7).

several programs), as well as the time periods and locations that will be included in the evaluation.

- **Design** defines the information that is needed to conduct the evaluation, data collection and analysis methods, and limitations to the evaluation.

Finally, analysts should assess the soundness (reliability and validity) of the evaluations they deem relevant to their fragmentation, overlap, and duplication review. Specifically, analysts should assess all selected evaluations against basic standards for research design, conduct, analysis, and reporting.¹⁸ Major weaknesses in these areas or limitations in the availability of program information that affect the reliability or validity of each study's findings must be identified and considered in using the study and placing confidence in the study's findings. Refer to Tip Sheet 3 in appendix III for questions that analysts can consider when assessing the soundness of identified evaluations.

3.2 Conduct a new evaluation if existing evaluations are not available, relevant, or sound.

If existing evaluations are not available, relevant, or sound, analysts should consider conducting a new evaluation to supplement their reviews. Executive branch leaders may also wish to conduct evaluations of their programs to demonstrate the results of and justify the continued funding for their programs, especially when resources are scarce. Designing an evaluation involves selecting appropriate measures and comparisons that will permit drawing valid conclusions about the programs. In 2012, GAO outlined key issues in planning evaluations of federal programs, including defining scope, selecting an evaluation design, addressing methodological challenges, and determining evaluation standards.¹⁹ Refer to Step 3.3 and Tip Sheet 4 in appendix III for resources on how to scope, design, and conduct an evaluation. When designing a new evaluation, analysts may wish to consider further assessing the positive and negative effects of fragmentation, overlap, or duplication identified in Step 2, in addition to program performance and cost-effectiveness.

3.3 Use evaluations to validate the actual effects of fragmentation, overlap, and duplication, and assess and compare the performance of programs.

Analysts can use results from existing evaluations that are relevant and sound or new evaluations to do the following:

- **Validate** the actual positive and negative effects of fragmentation, overlap, and duplication among the programs identified in Step 1 of this guide.
- **Assess** and compare the relative performance and cost-effectiveness of the identified programs.

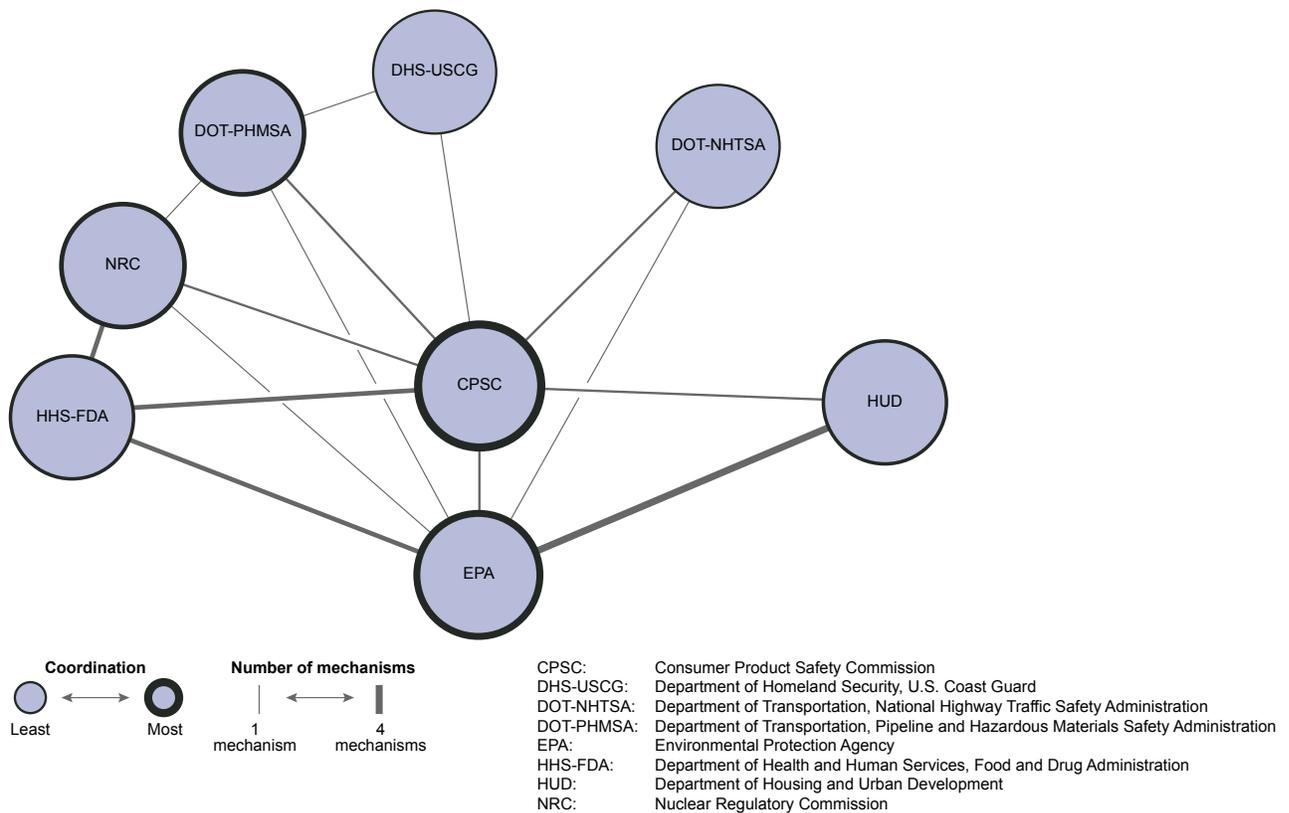
Validating the existence and extent of positive and negative effects can help guide analysts' recommendations on how to increase efficiency or reduce or better manage fragmentation, overlap, or duplication in Step 4. To illustrate, in its evaluation of consumer product safety oversight, GAO found that oversight of consumer product safety is fragmented across agencies, and identified instances in which jurisdiction overlaps or is unclear for certain products, and potential positive and negative effects of this fragmentation and overlap. In some cases, agencies regulate different components of or carry out different regulatory activities for the same product. In addition, regulatory jurisdiction for a product can change depending on where or how the product is used. To evaluate coordination among the agencies, GAO designed and administered a questionnaire to the identified agencies. GAO's evaluation also included interviews with agency officials and industry groups about the extent of fragmentation and overlap and their positive and negative effects and the extent of coordination among agencies.

¹⁸For example, see GAO, *The Evaluation Synthesis*, GAO/PEMD-10.1.2 (Washington, D.C.: March 1992). Analysts could also seek expert validation of the soundness of an evaluation.

¹⁹See GAO-12-208G.

Figure 4 illustrates the extent to which eight federal agencies coordinate on consumer product safety activities. In this figure, the thickness of the node rings represents the number of agencies that a given agency collaborated with and the thickness of the lines reflects the number of mechanisms that a given pair of agencies used to collaborate (which in this example includes memorandums of understanding or agreement, interagency agreements, working groups, and other practices, including coordination on an as-needed basis). The thicker nodes and lines signify more reported collaboration, while the thinner nodes and lines signify less reported collaboration.²⁰ For example, as shown by the thicker nodes in the figure, agencies most frequently coordinate with the Consumer Product Safety Commission (CPSC) and the Environmental Protection Agency (EPA). Similarly, EPA and HUD tended to use more mechanisms in coordinating with one another, as shown by the thicker lines in the figure than, for example, EPA and CPSC.

Figure 4: A Diagram of Coordination between Eight Regulatory Agencies on Consumer Product Safety Activities, as of November 2014



Source: GAO analysis of questionnaire responses. | GAO-15-49SP

The evaluation validated some positive effects of this oversight structure and coordination. For example, GAO found that agencies reported that the involvement of multiple agencies with various areas of expertise can help ensure more comprehensive oversight by leveraging each agency’s expertise to address a range of safety concerns, such as the National Highway Traffic Safety Administration and CPSC overlapping regulation of certain hand-held infant carriers that can also be used in cars.

The evaluation also validated some negative effects. GAO found that fragmentation and overlap created some inefficiencies and that jurisdictional boundaries are sometimes unclear. For example, in its evaluation, GAO found a lack of clarity regarding whether some recreational boating equipment should

²⁰For more information, see [GAO-15-52](#).

be regulated by the U.S. Coast Guard or CPSC. Coast Guard officials told GAO that they work informally with CPSC when the need arises, but that interaction is infrequent. GAO noted that without a more formal coordination mechanism to establish roles and responsibilities, some potentially unsafe products may go unregulated.²¹

Similarly, assessing and comparing the performance and cost-effectiveness of programs can help analysts determine which programs or aspects of programs are performing well and which are not, and by extension, which programs or aspects of programs to recommend for actions such as consolidation or elimination, as discussed in Step 4 of this guide. For example, the results from existing and new evaluations that have been appropriately designed can help analysts to:

- **Assess a single program:** How effectively and efficiently is the program performing? Does the program have the desired effect on the target population? What is the cost-per-good/service/product delivered by the program?
- **Compare two or more programs:** How do the performance and cost-effectiveness of programs compare?
- **Assess a system (or group) of programs:** How efficiently and effectively do these programs provide and support benefits, services, or products, individually and collectively? How would costs and benefits change if benefits, services, or products were delivered differently? Are there foregone benefits that could be achieved through integration or better coordination? How does the cost-per-good/service/product delivered by the system of programs compare to the cost-per-unit of success of the programs individually?
- **Evaluate interactions or relationships between programs:** Are relationships necessary, strong, and effective? Do they lead to better outcomes?

3.4 Confirm findings with relevant agencies and other key stakeholders.

Analysts should confirm their findings on the actual effects of fragmentation, overlap, and duplication, as well as program performance and cost-effectiveness, with relevant agencies and other key stakeholders. For example, analysts may wish to interview executive branch leaders and program administrators to obtain feedback on their analysis of existing evaluations. In addition, any limitations in the availability of program information identified during evaluations may help highlight areas where agency action is needed to develop sufficient evidence that analysts could use to guide their recommendations.

The results of this analysis should help solidify analysts' understanding of the effects of the identified fragmentation, overlap, or duplication and determine the relative performance and cost-effectiveness of the programs in providing benefits, services, or products. Analysts can use this information to help guide their recommendations to policymakers on how to increase efficiency or reduce or better manage fragmentation, overlap, or duplication, as discussed in Step 4.

²¹See GAO-15-52.