What Really Counts

Rethinking materiality as a practical, five-step process for more efficient, decision-relevant government reporting.

BY SHAYNE KAVANAGH AND MICHELE MARK LEVINE

Materiality plays a central role in financial reporting, yet it is often treated as a narrow technical judgment rather than an ongoing management discipline. When applied thoughtfully, materiality helps focus effort on information that matters to those who use the information. It also helps avoid effort wasted on low-value detailed accounting. Rethinking materiality doesn't entail relaxing standards or reducing accountability; it reduces low-value work, so finance professionals can focus on what matters to their communities.

This article explains how to apply materiality as a process. It is designed to help finance professionals identify opportunities to reduce low-value effort, assess qualitative and quantitative issues affecting materiality, and implement safeguards that preserve confidence in financial reporting. By treating materiality as a process rather than a single technical test, governments can better align reporting efforts with decision relevance. This strengthens both efficiency and accountability.

We outline a five-step process for rethinking materiality:

1. Identify the largest potential opportunities for reducing resource consumption and screen for associated risks.
2. Quantitatively evaluate the opportunities.
3. Confirm that identified risks can be sufficiently mitigated—that is, identify if and how qualitative factors should modify purely quantitative analysis.
4. Consider the cumulative impact of individually immaterial items.
5. Validate and communicate decisions to auditors and to internal and external stakeholders.

Step 1 Find the Opportunities

Step 1 uses the finance professional's intuition to find opportunities. The purpose of materiality judgments is to protect against decision-relevant risk. These risks include misleading conclusions about financial condition, results of operations, or loss of comparability on issues users demonstrably care about. This means that materiality is not about eliminating all error—it is about preventing decision-relevant error.

The search for opportunities starts from the premise that financial statements are the work product and responsibility of their issuers (the preparers), who are solely responsible for their content. Auditors are, of course, valuable resources and may qualify their independent audit opinions or even issue adverse opinions if they disagree with materiality or other financial reporting judgments. But the ultimate responsibility for all aspects of financial statements rests with the preparers.

Two questions guide our search for opportunities.

The first is whether you're putting more effort into measuring reported amounts than seems warranted by the benefit to the end user. Here, we rely on the preparer's intuition to identify opportunities, but that intuition will be tested quantitatively and qualitatively as we proceed. A critical assumption behind GFOA's Rethinking Materiality is that individual preparers can rely on the collective experience and intelligence of the public finance profession to make wiser materiality decisions. These learning opportunities should be designed by national- and state-level organizations, including GFOA.

A common opportunity is reportable activities where a high volume of small transactions does not make a difference in the user's understanding of financial condition. If an activity instantly comes to mind for the preparer, it's a candidate. Potential candidates include small

Rethinking materiality doesn't entail relaxing standards or reducing accountability; it reduces low-value work, so finance professionals can focus on what matters to their communities.

leases, minor software agreements, minor accruals, low-value or short-lived capitalizable items, and compensated absence liabilities for leave types used infrequently for short periods (such as jury duty).

Capitalizing on a savings opportunity does not necessarily involve ceasing to report the difficult-to-measure amount. Estimating a figure that was previously measured through detailed accounting is another option—we don't want to overlook three opportunities.

The second question is whether important qualitative barriers might stand in the way of realizing the opportunity, or whether qualitative considerations could lead to the conclusion that quantitatively insignificant amounts might nonetheless be material.

Qualitative considerations are essential to determining materiality. For potential opportunities, we must ask whether there are qualitative barriers to watch out for. Finding a barrier does not disqualify a topic from rethinking. Instead, consider strategies to overcome or bypass it. Let's examine four categories of barriers and bypass strategies.

The first category is **legislative** or **regulatory**. In this case, specific reporting may be required by local laws or regulations, by a higher level of government (such as federal or state governments), or by a grant, contract, covenant, or similar obligation. The first bypass strategy is to confirm that the "letter of the law" requires those details. Governments sometimes interpret regulatory language too conservatively.[1] If the language does, in fact, require the costly, detailed reporting, ask if an exception can be made—some regulators and grantors are willing to provide reasonable accommodations.[2] Finally, consider whether you can isolate the items that are subject to reporting requirements instead of applying the same accounting treatment to a broader category. For example, a grant agreement may require detailed reporting or specific accounting treatment for expenditures funded by that grant. Rather than applying those requirements to all expenditures within the broader category—such as all professional services or all capital-related costs—governments can often isolate the grant-funded portion and apply the required treatment only to those specific items.

The second category is **political** or **public sensitivity**, where a topic is likely to draw public attention. The primary bypass is like the grant example. See if the politically sensitive item can be "quarantined" from the larger category so that detailed accounting treatment doesn't need to be applied to the entire category. Also, it may be that politicians or the public do not care much about what accounting treatments are applied. The underlying concern may really be a budget or policy issue that could be addressed in some other way. For example, public safety overtime or travel expenses are often a topic of heightened public and political interest. In financial reporting, these costs may be embedded within a much larger public safety function. A separate analysis and presentation of overtime spending and trends outside of annual financial reporting would likely be the best way to approach this topic.

The next category is **auditor attention**. First, recall that materiality determinations are primarily the responsibility of financial statement preparers—that is, the issuing governments. The auditor's role is to review the preparers' judgments and decide whether they are reasonable. If auditors are asked first, their interests generally align with having preparers make conservative materiality determinations (that is, inclusive of more items), and preparers may feel pressure to defer to the auditors' judgments. The strategy here is for the preparers to first develop and document management's professional judgment and then address it with the auditors.

Our final barrier is **control dependence**, which refers to the risk of harming a control that depends on the existing approach to materiality. For example, imagine a grant that requires tracking the assets that the grant funds. These assets cost as little as $10,000. Keeping the $10,000 threshold for all assets can trap the government in a great deal of low-value accounting work.

So, you've decided to rethink materiality...

In 2025, GFOA published "Finding the Balance: Rethinking Materiality in Government Accounting," which made the case for the profession to move beyond a compliance-oriented view of materiality toward a more judgment-based, value-centered approach. That report reframed materiality as a professional determination about what information is likely to influence user decisions. This article builds on that foundation by showing how to turn the intent to rethink materiality into action.

READ THE REPORT: gfoa.org/materials/finding-the-balance-rethinking-materiality

In this case, we could ask whether there is a way to satisfy the terms of the grant without also doing the accounting work for potentially hundreds of additional assets that do not affect the judgments of the financial statement users.

The next step is to test our intuitions to see if the opportunities are real.

Step 2 Quantify the Opportunities

Now that we know where the opportunities are and which qualitative risks matter, Step 2 helps us quantify those opportunities. Use the data to confirm whether the opportunities are real. If so, we can use the data to determine how materiality should be applied. To accomplish this, it's helpful to distinguish between three related but distinct decisions.

The first is a "go/no-go" decision. Here, we are asking if an entire activity should be reported at all. For example, a government might not have a large enough compensated absence liability to warrant reporting.

The second is a "reporting threshold" decision. This determines which individual items within a reportable activity are large enough to be included in the report and which aren't. A common example is a capitalization threshold for capital assets.

The reporting threshold determines what is included. The precision threshold determines how much work goes into measuring what is included.

The third decision has not previously been explicit in public finance and is what we will call a "precision threshold" decision. This determines how much measurement effort is warranted once an item is deemed necessary to include in the report. Amounts below the precision threshold may be reported using reasonable estimates rather than detailed transaction-level accounting. The reporting threshold determines what is included. The precision threshold determines how much work goes into measuring what is included.

These three decisions serve different purposes. The first two govern inclusion. The third governs effort. All three work together to reach materiality decisions that are savvy, defensible, and contribute to a high-value financial reporting process.

We'll now go deeper into how to approach these three decisions.

Applying go/no-go decisions

A go/no-go decision encompasses an entire category, such as whether it's necessary to report compensated absence liabilities. Other examples of categories include pensions or other postemployment benefits (OPEB), lease and other long-term liabilities, and right-to-use lease, subscription, and public-private and public-public partnership (PPP) assets and their related liabilities. We start with the preparer's intuition that reporting a given activity does not add value to the financial report, but we don't end there. We test this intuition by seeing if excluding the activity moves key financial health ratios and other appropriate indicators, such as debt margins. If the ratios remain stable, that supports the presumption of immateriality, while recognizing that any qualitative landmines still need to be addressed.

Exhibit 1 provides examples of activities that could be candidates for a go/no-go decision, why no-go might be reasonable, and the ratios that support or rebut the presumption of immateriality.

Use Over-Precision as a Clue for Finding Opportunities

Think about where being more precise might be causing more work than it's worth. An example is the discretionary categorization of one or more governmental funds or enterprise funds. The Governmental Accounting Standards Board requires that funds meeting certain relative size thresholds be reported as major funds, but it permits preparers to report additional funds at their discretion. Independent auditors must provide opinions on each individual major fund. This means that quantitative materiality determinations are based on smaller individual fund bases—such as total assets and total revenue—for each major fund. For nonmajor funds, materiality is calculated using the aggregated bases of all nonmajor funds. Governments are therefore much more likely to adjust financial statements for issues identified in a discretionary major fund, even though the amount would be deemed immaterial if the fund were reported as nonmajor. Governments should therefore ensure that decisions to discretionarily report certain funds as major are based on substantive user needs rather than routine practice.

Our suggested ratios cover two factors: the size of the category compared to the larger overarching category users are likely to care about, and the potential for rapid or sudden changes in the reported amount.

If the preparer has been doing the detailed accounting work for the activity in the past, it's easy to use prior-year results to conduct our suggested ratio checks; however, there may be cases where historical data is not available. Perhaps the preparer is implementing a new GAAP pronouncement or the activity was deemed immaterial for several years, but changes in circumstances require a reassessment. In such cases, consider whether estimates can be used to get a sense of the impact. The GFOA paper, "Close Counts: Expanding the Use of Estimates for Better, Faster Financial Reporting," can help and be read at gfoa.org/materials/close-counts.

As you consider potential no-go's, recall two things. First, every Statement of the Governmental Accounting Standards Board contains a "materiality box" at the end of the authoritative guidance section that states, "The provisions of this Statement need not be applied to immaterial items." Regardless of the exact language used, the standards are expressly applicable only to material items.

After confirming that an activity is a "no-go," distinguish between a reporting approach that would risk negligence and one that applies reasonable limits to what will be reported while remaining vigilant. One effective technique is to set tripwires that trigger a review of the decision to treat an activity as immaterial.

For example, with compensated absences, a tripwire could be a change in personnel policy that provides more generous leave benefits or creates many new positions that are eligible for defined benefit post-employment benefits. Either change would trigger a re-evaluation of the earlier decision.

If an activity is designated as a "go" but the preparer still believes the reporting effort is not justified, then savvy decisions about the reporting and precision thresholds hold the key to realizing the opportunity the preparer senses.

EXHIBIT 1 | Potential Go/No-Go Categories and Key Ratios for Assessing Immateriality

Category	Indicators or Ratios to Check
Compensated absences	▪ Accrued leave liability as a percent of total liabilities ▪ High-growth or volatility trends
Postemployment benefit (PEB) pensions and other post-employment benefits (OPEB)	▪ Net pension or OPEB liability as a percent of covered payroll or total liabilities ▪ Volatility in PEB estimates or discount rate sensitivity
Lease, subscription, and PPP liabilities	▪ Lease or PPP liability as a percent of long-term liabilities ▪ Change in lease activity from prior year
Right-to-use lease, subscription, and PPP assets	▪ Percentage of capital assets comprising each category

Using the shape of the data to guide threshold choices

To make threshold decisions in a principled, defensible way, it helps to understand the underlying "shape" of the relevant data distributions. Shapes tell us where value and volume are concentrated and, thereby, how thresholds should be applied. We focus on two shapes: the bell curve and the Pareto pattern. Both appear frequently in data and provide strong, intuitive signals for judging materiality. Further, these two shapes are derived from widely accepted and powerful statistical principles.

Our goal is not to turn accountants into statisticians. Rather, we will borrow the concepts behind these shapes from statistics and show how they can support savvy thinking about materiality without necessarily requiring statistical calculations.

Bell-shaped patterns: modest reporting threshold gains, big precision threshold gains

This is a shape that is common in many data sets and familiar to many readers: the bell curve. A bell curve describes a situation in which the "average" item is the most commonly occurring type of item. Relatively few items are much larger or smaller than that average, and those larger and smaller items are evenly distributed around it.

A familiar example of a bell curve in financial reporting is pension investment returns when viewed across many years. Over time, most returns tend to cluster around the long-term average, with fewer returns deviating greatly from the average (e.g., particularly good or bad years). A more useful example for materiality decisions is payroll-related accruals, which often follow a bell-shaped pattern. Small timing differences in hours worked, pay cycles, or payroll cutoffs tend to balance out across hundreds or thousands of employees. Errors in one direction are usually offset by errors in the other.

Exhibit 2 shows a bell-shaped distribution for year-end utility bills, illustrating how public utilities might estimate customer receivables at fiscal year-end. In real life, data sets may not be perfectly symmetrical around the average, but the basic idea still holds: most items cluster near the average, differences from that average are evenly distributed, and extreme values are rare.

The bell curve holds two clear implications for thresholds.

First, because very small items are relatively infrequent in a bell-shaped

EXHIBIT 2 | The Bell Curve

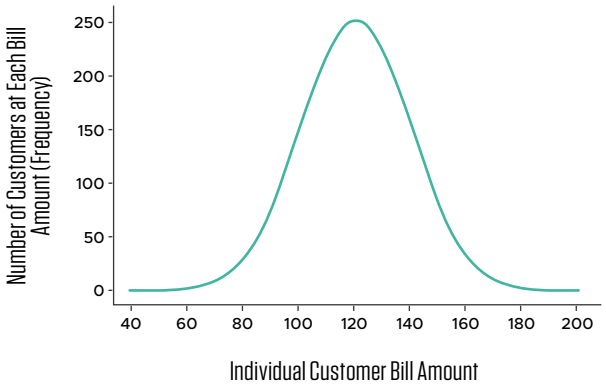
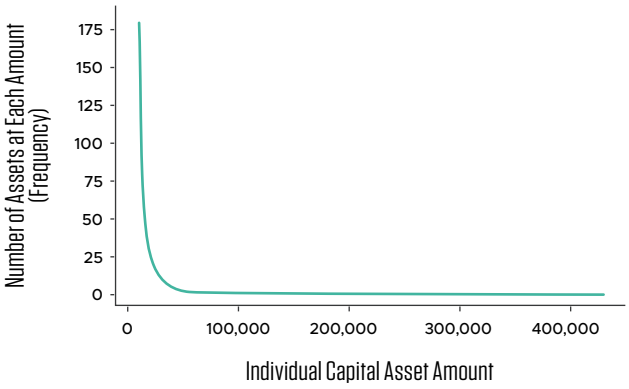


EXHIBIT 3 | Pareto Distribution



pattern, there's usually less opportunity to raise the reporting threshold without excluding a nontrivial share of the total dollar value.

Second, the predictability of the bell curve creates a significant opportunity to raise the precision threshold. When most items cluster near the average and deviations are balanced, a well-designed estimate can capture the total with reasonable accuracy without transaction-by-transaction accounting.

In other words, bell-shaped categories are often better candidates for estimation compared to excluding large swaths of items from reporting altogether. The savings come from doing less detailed work on items already included in the report, not from dropping them entirely.

Back-Testing: Validate Your Estimates

Estimates can be tested by applying them to older data sets where you have done the detailed accounting to see if your estimate is sufficiently close to the actual result.

Below are several examples of topics that could fit this category. Each government will need to evaluate the potential of any given opportunity for itself, including how those opportunities might change from year to year, given evolving circumstances.

- Accrued compensated absences (unused vacation or sick leave).
- Routine small accounts payable accruals (supplies and services).
- Purchasing card (P-card) purchases at year-end.
- Receivables for utility charges.

Pareto patterns: strong reporting threshold and precision threshold leverage

The second shape is the Pareto pattern, which can be powerful for making materiality decisions. It is characterized by a small number of very large items and a larger number of small items. The total value of the large items is often overwhelmingly greater than that of the small items. The Pareto distribution is the source of the well-known 80/20 rule, or Pareto Principle, which describes cases in which about 20 percent of the causes account for about 80 percent of the effect. Applied to financial reporting, we often find cases where about 20 percent (or less) of the items in a data set account for 80 percent (or more) of its total value.

Examples of data sets that often follow a Pareto pattern include capital assets, leases, and certain types of accruals. Exhibit 3 illustrates a Pareto distribution for capital assets. Notice that the most frequently occurring quantity is now on the far left side, not the middle. The tail also extends much further to the right. These are critical differences from the bell curve.

In a Pareto setting, the reporting threshold can do much more work. Because much of the total value sits in a small number of large items, the reporting threshold can be raised significantly—removing large numbers of low-value items from detailed accounting—while changing the total reported value by only a small amount. A precision threshold can further reduce effort for mid-sized items by allowing reasonable estimates below that threshold. Exhibit 4 is a different way to show the Pareto pattern, called a Lorenz curve, using real data from the City of Oak Harbor, Washington. It shows how concentrated the values in this data set are. The horizontal axis represents the cumulative share of all items, from smallest to largest. The vertical axis shows the cumulative

share of the total value those items account for. If every item contributed equally, the curve would follow the 45-degree line of equality. The more the curve bows below that line, the more unequal the contributions are, meaning a small number of items account for most of the total. In other words, a steep Lorenz curve indicates strong Pareto behavior: a few large values dominate the outcome. The exhibit also shows two dots representing the current and proposed new threshold. We will see why this is useful shortly.

The Pareto pattern tells us we can pick a reporting threshold larger than the individual value of many small items without changing the total reporting value by much. Consider the city's current reporting threshold of $5,000. Using this threshold, the city accounts for 901 asset records. Exhibit 5 shows potential new thresholds, how many assets the city could stop reporting if it moved to the new threshold, and the total change in the reported value of the city's assets if the new threshold were put in place.[3]

We see that it's possible to stop reporting a large number of assets with little change in reported value. For some of the thresholds in the table, the city could recover a non-trivial number of staff hours spent accounting for assets in exchange for a trivial change in reported value. Going back to Exhibit 4, if we add a proposed threshold of $50,000 (blue dot), which is 10 times the current threshold (yellow dot), the blue dot is barely higher on the vertical axis than the yellow dot, showing that the cumulative share of value hardly changes.

The real power of the Pareto pattern comes from considering the reporting and precision thresholds together. Suppose the city sets the reporting threshold at $10,000, so it no longer reports assets below that amount. The new reported value would be $332,414,319. It might then place its precision threshold at $50,000, eliminating detailed accounting work for assets below that level. This precision threshold would not appreciably change the total reported value from $332,414,319. This is because the city would estimate the value of assets between $10,000 and $50,000, as well as their related annual depreciation expense. The estimate may not be exact, but it would be close enough to avoid negatively affecting the decision value of the statements to the user.

Let's do the math. The true value of the city's assets above $10,000 is $332,414,319. The true value above $50,000 is $323,357,236. The city is therefore estimating the difference between these two figures, or about $9 million. The city could be off by 33 percent when estimating the value of assets between $10,000 and $50,000. Even with that error, the reported value of all assets above $10,000 would still be about 99 percent accurate. Of course, the city can do better than a 33 percent error on its estimate, but in a Pareto distribution, accounting precisely for the fewer large items leaves a lot of room for less precise estimates among the more numerous but smaller items.

Preparers should consider the shapes of data sets for the opportunities they have identified. There are simple checks to assess the shape. Start by sorting the data by item size. If the top 20 percent of items account for about 80 percent or more of the total value, the data follows a Pareto pattern.[4] If most items cluster around the average value with relatively few outliers, the data is bell-shaped.

Pareto pattern opportunities are usually easier to realize. Preparers start there and keep bell-shaped and other opportunities in reserve. After gaining experience rethinking materiality through the Pareto pattern, they can then turn to those reserved areas. GFOA's "Close Counts: Expanding the Use of Estimates for Better, Faster Financial Reporting" offers more guidance on the Pareto pattern and its use in estimating, which is key to raising the precision threshold successfully.

EXHIBIT 4 | Concentration of Values in the City of Oak Harbor's Capital Assets

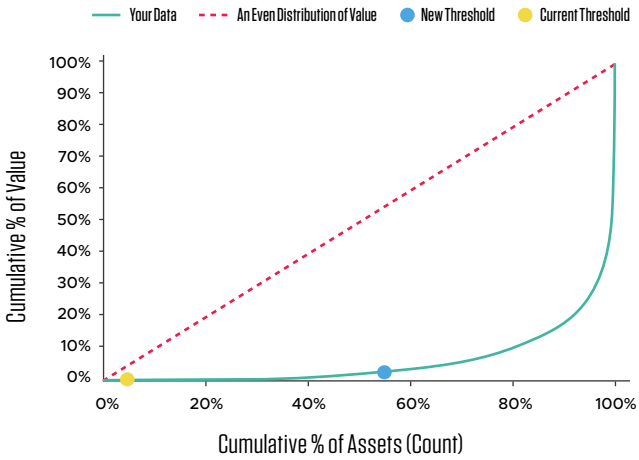
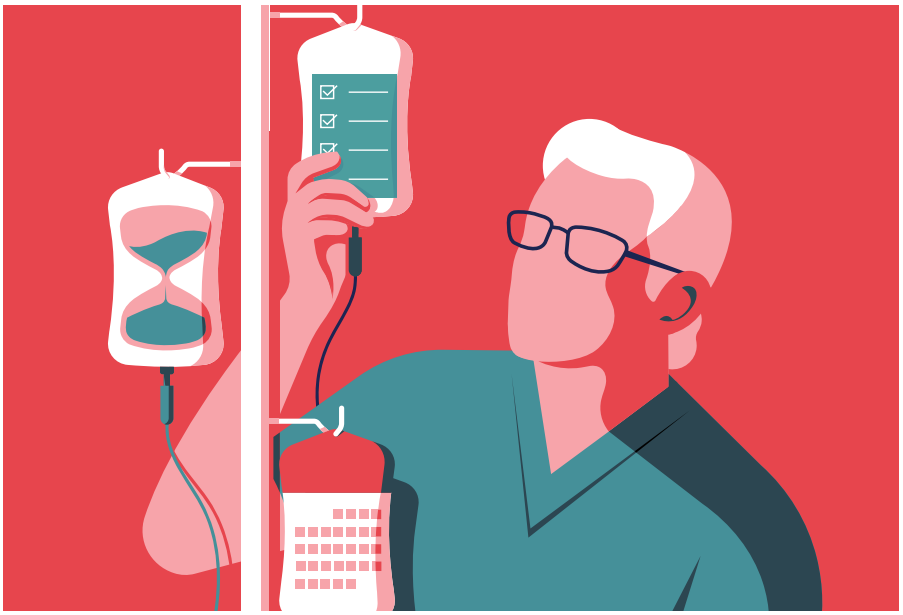


EXHIBIT 5 | Potential New Thresholds

New Reporting Threshold	Reduction in Asset Records (of the 901)	% Reduction in Reported Assets	% Reduction in Total Reported Value
$10,000	126	14%	0.30%
$20,000	277	31%	0.97%
$30,000	365	41%	1.64%
$40,000	424	47%	2.26%



Step 3 Confirm Risks are Mitigated

If Step 2 shows potential from a different approach to materiality, Step 3 proves the mitigations hold in practice. This involves stress testing solutions to the barriers identified in Step 1.

For legislative or regulatory risks, we ask whether the new approach to materiality is consistent with the regulation as written. The finance professional could share an example of the new approach with regulators to see if they object. The profession, as a group, could also lobby for federal uniform guidance to allow the use of statistically valid estimates for grant reporting and compliance.

For political or public risks, we can offer an example of an alternative approach to communicating or reporting the information that addresses the political concern at a lower cost. For example, transparency around public safety overtime or travel costs could instead be provided through a website that allows users to make custom inquiries of financial data.

For risks related to auditor concern, share the strategy with auditors as early as possible in audit planning to learn if they will sign off.

For control risks, develop a compensating control and test whether it is fit for purpose. To continue our example of the grant requirement, see whether the information required by the grantor can be assembled without extending detailed accounting treatments to a much larger category of items that aren't subject to the requirement.

Step 4 Test Cumulative Impact

In Step 4, we consider whether efforts to rethink materiality could die of a thousand immaterial cuts—that is, whether many small exclusions or estimates could accumulate into something that matters for decision making.

First, we must recognize that the risk of cumulative impact is slightly different for activities deemed a "no-go" (i.e., won't be reported at all) versus activities that are reported but where we have applied new reporting and/or precision thresholds.

Low-cost, high-assurance checks can verify that the cumulative impact remains within acceptable limits.

Let's start with activities where new thresholds have been applied. For most governments—and data shapes—the risk of a material cumulative impact is negligible. Low-cost, high-assurance checks can verify that the cumulative impact remains within acceptable limits. First, we'll review why cumulative impact isn't likely to be a problem and then show low-cost ways to confirm it.

To start, cumulative materiality does not mean that each year's differences are added to the prior year's differences. In most cases, those differences will self-correct over time. For example, if too much asset value is expensed each year rather than capitalized, the "understatement" of depreciation expense in subsequent years would offset the difference. As a result, it would exist only for the useful life of the uncapitalized assets. Moreover, expensing of newly acquired uncapitalized assets each year would likely result in a relatively consistent difference year to year, unless the government's small asset acquisition pattern is notably irregular.

The main reason cumulative impact is unlikely is the data- and shape-informed system introduced in Step 2.

First, each threshold is set relative to the shape and scale of the data that make up the activity (e.g., capital assets and leases). That means any individual threshold is sound in the context of its activity. Individually sound thresholds are unlikely to add up to collectively unsound results.[5]

Second, the dual-threshold system in Step 2 further reduces cumulative risk. If the reporting threshold were the only way to rethink materiality, there could be a problem. It is conceivable that amounts falling below reporting thresholds across activities could add up if the thresholds were set too high. The precision threshold provides the solution. Instead of excluding transactions entirely, we estimate them and avoid detailed, low-value accounting work. This doesn't mean the reporting threshold should never be raised. It means the precision threshold can help avoid the risk of raising it too much while still benefiting from rethinking materiality.

The skeptic may raise a concern that potential errors across numerous estimates might add up to an amount large enough to negatively impact the decision usefulness of the report. The good news is that research confirms that aggregation often works in favor of the preparer.[6] There are three features of estimates that contribute to stable, predictable results.[7]

Unbiased estimates avoid error stacking. Researchers have found that private firms seeking to produce impressive financial results sometimes exhibit optimism bias in their estimates. This bias can stack up across different activities, creating a meaningful cumulative impact.[8] Governments don't have the same incentives as private firms, but bias can still occur. For example, conservative estimates may seem prudent—but conservative bias also stacks up. The solution is simple. Use your best estimate of the actual value, without hedging in either direction.

Numerous estimates offset errors. When estimating many smaller items, unbiased errors tend to offset one another.[9] In Step 2, we emphasized that precision thresholds should apply to the numerous smaller items in an activity, not the larger ones.

Uncorrelated estimate errors avoid error stacking. Errors in different estimates are unrelated to one another. To illustrate, suppose item A is overestimated. If knowing that item A is overestimated allows us to predict that item B will also be overestimated, the errors are correlated. This is a problem because the errors will never balance out (thus defeating the advantage of numerous estimates). We want uncorrelated errors, where an error in item A tells us nothing about the error in any other item. We have already addressed the largest source of correlated risk: biased estimates. Using our best estimate of the actual value will go a long way toward addressing this problem.

Another source of potential correlated error is if a single, blunt estimation method is applied across several different activities. In theory, if a government applied a flawed method broadly, it could create widespread, consistent errors. For example, applying a single generic growth rate to all historical amounts to estimate current values could have this effect. In practice, this is unlikely. Because governments report many kinds of transactions—like capital assets, accruals, and compensated absences—a one-size-fits-all method would not be used. To avoid this problem, the solution is: Do not apply overly generic estimation methods across multiple activities.

Finally, we should highlight important consistencies between what we just covered and the shape-based approach in Step 2. In bell-shaped categories, individual items cluster around the average, and deviations in both directions tend to cancel out when aggregated. This makes the total predictable and limits cumulative drift. In Pareto-shaped categories, most of the total value lies above the precision threshold by design. The numerous items below the threshold are less likely to collectively produce large errors.

To address any remaining doubts about the likelihood of materially incorrect cumulative results, we can use common techniques to check cumulative impact without recreating full accounting detail.

One such technique is the **envelope test**, a form of stress test used to assess whether doing less than the full accounting for all transactions in an activity could negatively impact decision usefulness. The test assumes that conservative estimates are closer to the truth and asks whether that assumption would change user judgments.

To check the reporting threshold:

1. Estimate the number of items below the threshold.
2. Assume each item equals the threshold value, meaning each excluded item is as large as possible without exceeding the threshold.
3. Determine whether this hypothetical total would change cumulative amounts enough to negatively impact decision usefulness.

To check the precision threshold, use appreciably more conservative assumptions in the estimate. For example, if the estimate uses an inflation rate, apply a more conservative rate. Check to see if the new estimate would change cumulative results enough to raise concern. This technique is also known as "sensitivity analysis."

Finally, you can apply the test to a single activity where there is a particular concern that threshold choices are contributing to a cumulative problem. Be very careful about applying the test to multiple activities simultaneously; doing so would create the correlated conservative bias risk described earlier. Also, the chances that highly conservative assumptions are closest to the true results across several activities at the same time are small, so the test becomes highly unrealistic in such circumstances.

Rotation sampling or spot checking is a way to validate assumptions related to threshold choices that could impact cumulative results. To do this, select a random sample of 10 to 20 items below the threshold and measure their actual value to see if the results are consistent with the assumptions. This test is particularly useful for precision threshold

estimates because you can compare the average value of sampled items to the assumed average. It can also be useful for reporting thresholds. If too many items are close to the threshold value, it would suggest that the conservative envelope test is, in fact, descriptive of reality.

Comparing the estimate to actuals is a niche test useful for a group of related items, such as expense and revenue accruals and collectability assumptions, where actuals become known after the report is issued. If the test shows the difference is minimal, the method is sound. If the difference is large enough to compromise the decision usefulness of the numbers, then the method should be revised.

Together, these checks provide proportionate reassurance that small omissions or estimates do not accumulate into a material impact. If these tests suggest that cumulative effects could plausibly influence user judgments—or if the area is particularly complex or politically sensitive—more sophisticated estimation or validation tools may be appropriate. Even then, standardized templates or applications can provide this capability at far lower cost than full re-performance of detailed accounting work.

Materiality is not about eliminating all error—it is about preventing decision-relevant error.

Now, let's shift our focus to the risk of cumulative impact arising from any activities that have been excluded entirely from reporting (the no-go's). The concern is if a government were to deem multiple activities entirely immaterial, those activities might have an impact when considered together. For example, perhaps leases and compensated absences are both deemed immaterial, individually. If they were added together, might the impact on total liabilities matter?

The first step is to test the proposition by estimating the impact of those two activities. The GFOA paper, "Close Counts," provides insights into using estimates for financial reporting.

Let's suppose the test suggests that there could be a material cumulative impact. The solution is not necessarily to do the detailed accounting work for the no-go's. The most expedient solution could be to use the estimates in the report. For activities that were too small to include individually, the error in the estimate would have to be huge to have a material impact. As a simple example, let's imagine you want to be 99 percent accurate in your reported liabilities. It turns out that compensated absences comprise 1 percent of all liabilities, and leases comprise 1 percent. So, if you excluded both, you'd only be 98 percent accurate. Your estimate of the combined liability from both leases and compensated absences could be as large as 50 percent, and you would still be 99 percent accurate. Of course, preparers can do much better than a 50 percent error.

If, for some reason, it isn't possible to report the estimate, another option would be doing the accounting for the smallest number of excluded activities to get to your preferred accuracy level. In our example above, reporting either leases or compensated absences would be sufficient for 99 percent accuracy. It would not be necessary to do both.

Finally, we'll address fund-level versus government-wide materiality considerations. First, if prudent estimation methods are applied to any category of financial information, you will likely be materially accurate for that category. If each component of a

total (government-wide or fund level) is materially accurate, the total will be as well. Second, think of materiality in proportionate, not absolute, terms. For example, $100,000 might be trivial relative to a total reported value of $100 million but nontrivial relative to $1 million.

Step 5 Validate and Communicate

Once thresholds are set and the lack of cumulative impact is confirmed, the final step is to validate the decisions internally and communicate them externally. The goal is to show that materiality judgments are well-founded, risk-aware, and subject to appropriate oversight, so they can withstand scrutiny and deliver durable value.

Documentation: make the logic visible

Maintain a clear internal record that describes:

- How reporting and precision thresholds were set and why.
- How cumulative impact was checked.
- How financial statement elements are affected.
- How the process will be monitored over time.

This will help auditors, colleagues, and successors follow the rationale behind materiality judgments. There might also be value in including notes in financial documents to clarify those decisions. Auditors for the City of Newport Beach, California, included the following in their required communications with the Newport Beach City Council: "The city applies a capitalization threshold for lease receivables consistent with its accounting policies. As a result, certain lease receivables below the threshold—totaling approximately $1.86 million—were not recorded. These amounts are not considered material to the financial statements."

Tripwires: make materiality decisions adaptable

Materiality decisions are not "set-and-forget." They must adapt over time.

That's where tripwires come in: predefined events or changes that automatically trigger a review of materiality decisions. These might include:

- A major policy or service change that alters transaction patterns. For instance, a natural catastrophe may cause a sharp increase in P-card purchases for emergency materials. In that case, year-end P-card purchases should not be estimated.
- Major new grants (e.g., COVID-related funding received by states and local governments), changes in the composition of major funds, or updates to generally accepted accounting principles, including new GASB pronouncements.
- Accounting system changes that make measurement more or less demanding.

Tripwires help ensure that once-reasonable materiality decisions don't linger past their useful lives.

Auditor engagement: prevent surprises

When external auditors understand the logic and safeguards behind materiality decisions, they're more likely to support the result. Share the following:

- The distribution-shape-based framework used to set thresholds.
- The tests used to check cumulative impact.
- The tripwires that prompt future reevaluation.

This builds confidence that materiality judgments are part of a savvy, risk-aware, value-centered financial reporting approach.

Remember that rethinking materiality focuses on providing decision-accurate information to the public as quickly as possible and at the lowest cost. It does not represent a lowering of standards. On the contrary, it represents maximizing the value of public resources.

Where to Go from Here

Rethinking materiality should not be done by individual preparers acting alone—local finance professionals will be far more successful if there's an overarching "permission structure" for rethinking materiality. A community forum for discussing materiality will increase the rate of progress, as well as lead to wiser decisions about how to produce the financial information the public needs at an acceptable cost.

Shayne Kavanagh *is the senior manager of research at GFOA's Research and Consulting Center.* ***Michele Mark Levine*** *is the director of GFOA's Technical Services Center.*

[1] Frederick M. Hess, *Cage-Busting Leadership* (Harvard Education Press:2013). Hess examines cases in which state education agencies offered to waive or modify state rules that local districts believed were hampering effectiveness. In most cases, waivers were unnecessary because existing rules did not prohibit districts from taking their desired course of action.

[2] Ibid.

[3] In practice, governments increase capitalization thresholds prospectively only; however, an analysis such as this—showing that differences would have been immaterial had changes been made earlier—provides strong support for making such a change.

[4] Technically, if the concentration of value in a small number of items exceeds what the 80/20 rule suggests, the data might not be "Pareto"; however, these more extreme distributions are even stronger candidates for the kinds of treatments we've suggested for Pareto data sets.

[5] This principle is supported by the central limit theorem, which shows that when individually small, independent deviations are aggregated, their cumulative effect tends to converge toward a predictable average, limiting the risk of material cumulative misstatement.

[6] General forecasting science has repeatedly shown that combining forecasts improves accuracy. See: J. Scott Armstrong, *Principles of Forecasting* (Springer Science & Business Media: 2001). We combine forecasts by combining estimates of many small items. More directly related to accounting, Arya and Glover discuss the cancellation of errors through aggregation as part of a broader discussion on the benefits of aggregation in accounting. See: Anil Arya, and Jonathan C. Glover, "Aggregation and the measurement of performance," *Journal of Management Accounting Research*, 26 (2), 2014.

[7] These conclusions are consistent with widely recognized statistical principles like the central limit theorem and variance sum law.

[8] Timothy A. Seidel, Chad A. Simon, and Nathaniel Stephens, "Management bias across multiple accounting estimates," *Review of Accounting Studies*, 25, 2020.

[9] This reflects the central limit theorem (CLT), which states that when a large number of independent, unbiased random variables are aggregated, their sum tends toward a normal distribution, and individual over- and underestimates tend to cancel out. George Casella and Roger L. Berger, *Statistical inference* (Thomson Learning: 2002).