What Counts and What Is Counted:
Seeing Organizations Through an Accountant’s Eyes

Prof. Robert Bloomfield
Johnson Graduate School of Management
August, 2012
ABOUT THE COVER

The picture cover photograph represents a flower as seen through the eyes of a human (left) and a bee (right). Flowers that appear similar to us look very different to a bee, which can see colors in the ultraviolet range of the spectrum. In the same way, situations within an organization that appear similar to the typical manager look very different to the accountant, who can see distinctions described by the rich set of terms and techniques described in this volume.
Part I: New Eyes for the Right Nose
These essays are required reading for my MBA and Executive MBA courses in Managerial and Cost Accounting at Cornell University’s Johnson Graduate School of Management. Topics range far beyond accounting, because my goal is to provide a comprehensive framework for understanding how reporting systems help organizations respond to their challenges. Many challenges arise from organizations’ interactions with outside parties, so some essays address product markets, supply chains, and investor goals. Others arise within the organization itself, requiring careful analysis of how people in the organization interact with one another, how employees’ interests may diverge from the interests of the organization, and how the interests of one group of employees may differ from those of another group.

The title of this collection is inspired by two quotations. The first is from a sign said to hang in Albert Einstein’s office (though according to this link should be attributed to sociologist William Bruce Cameron):

“Not everything that counts can be counted, and not everything that can be counted counts.”

A constant challenge in managerial reporting is to distinguish clearly between what truly counts to an organization, and what the organization actually counts. For example, an organization might care deeply about customer satisfaction; that is what counts. But they may only be counting the number of customers who report being “mostly satisfied” or better in a Web-based survey that only 10% of customers fill out. What counts (the true state of customer satisfaction) has long-run implications for the organization that are hard to observe, but what the organization counts (the answers to the survey) has an immediate tangible effect on the marketing director’s compensation and tactics.

The difference between what counts and what is counted is usually clear once the distinction has been pointed out. But too rarely is it pointed out! Even professional accountants often forget to distinguish between the true increase in an organization’s wealth over the year (what counts) and the net income number reported at the bottom of the organization’s income statement (what is counted). This observation—that people often miss what becomes obvious once it is pointed out—motivates the second quotation alluded to in the title, attributed to French author Marcel Proust:

“The real voyage of discovery consists not in seeking new landscapes but in having new eyes.”
Introduction and FAQs

My goal is help readers see the organization and its reporting needs through new eyes—the eyes of an accountant. The landscape still consists of familiar organizations and the common problems they must address. But the essays provide you with a rich language you can use to make careful distinctions among the many challenges organizations face, the solutions reporting systems provide, and the causes and consequences of the mismatches between challenges and solutions that arise all too commonly.

The remainder of this essay poses and answers some Frequently Asked Questions to clarify what material we will cover, how I approach it, what will be expected of you, and why.

What Is Management Accounting, and Why Should I Study It?

Management accounting is the study of how to (1) design the management reporting systems an organization needs to execute its strategy effectively, (2) interpret the outputs of those systems to evaluate people, products, projects, and processes, and (3) propose changes to both the systems and the strategy when circumstances warrant.

To do this job well, you need to answer questions such as the following:

- How can I define my strategy precisely enough to know whether I am executing it successfully?
- What information do my managers need to make intelligent decisions?
- Why types of management reporting systems give managers that information?
- How can I encourage my managers to pursue the organization’s goals, not just their own?
- How do I determine which managers or products are responsible for good or bad performance?
- How do changes in my marketplace affect how I should manage my organization?

Regardless of your industry or functional area, you will benefit from being able to answer these questions, and this course gives you the tools to do so.

What If I Don’t Want to Be an Accountant?

Very few people aspire to be accountants. But even if you don’t want to become one, you will definitely work with them—and quite likely, work for them. Despite what you may believe about their personalities, accountants are among the most likely people to become leaders of their organizations. Approximately a quarter of Fortune 500 CEOs have spent time in an accounting-related function. Virtually all of the CEOs in the UK’s FTSE 100 (the London Stock Exchange’s index of 100 large firms) hail from accounting. Accountants are strongly represented on boards of directors and in executive roles just below the CEO level. You can see a similar pattern here at the Johnson School. Over the last 20 years, accountants have held a disproportionate number of leadership positions, including dean, acting dean, associate dean, and director of the doctoral program. Accountants move up through the ranks not because they know how to interpret financial reports (though that helps), but because they understand the design of organizational systems and how people respond to those systems. One or two courses in accounting won’t make you an accountant, but they will help you understand how the accountants you work with—and for—see the world.
What If I’m Already an Accountant?

Even seasoned professional accountants will encounter a good deal of new material in this course, and learn to see familiar material through new eyes. Unlike most accounting courses, we devote as much attention to the needs of the organization as to how to meet those needs through accounting and reporting systems. What is your competitive position? How predictable is your production process? How variable is demand? How much risk in pay are your workers willing to tolerate? The answers to these questions determine what types of reporting systems most effectively serve your organization. Also, most accounting courses focus heavily on the rules accountants use to measure costs and income, giving students (and professional accountants) the impression that the number labeled “Income” or “Cost of Goods Sold” on a financial statement is the true income or cost of goods sold. In contrast, this course emphasizes that the numbers on financial statements are merely imperfect reflections of what we are trying to measure; I devote much of the course to understanding those imperfections and explaining why they matter.

What Background Do I Need to Succeed in the Course?

Management Accounting is one of the most interdisciplinary courses in the Johnson School. We touch on issues addressed in all of the core courses (Accounting; Economics; Finance; Marketing; Operations Management; Quantitative Methods, Management, Leadership and Organizations; and Strategy), as well as a number of electives (including Financial Statement Analysis; Marketing Strategy; Negotiations; and Supply Chain Management). Students who are familiar with this material will naturally be more comfortable in the course. However, I introduce each topic from scratch, so if there is material you haven’t seen or don’t remember as well as you’d like, you will have the time and materials you need to get up to speed.

If you are uncomfortable with math and bookkeeping (the technical side of accounting), don’t be too worried. We never use math that goes beyond middle-school algebra. You need to become comfortable with the basics of double-entry bookkeeping (debits and credits), but again, I teach this from scratch, and provide many tutorials (written and video) to help you. If you aren’t mathematically or technically inclined, this part of the course will be about as challenging as the average accounting topic covered in the core curriculum (easier than the hardest ones, but harder than the easiest ones). But technical topics constitute only about a third of the material we cover. The majority of the course focuses on teaching students the terminology needed to understand management accounting.

Why Focus on Terminology?

The backbone of these essays is a list of hundreds of terms, carefully defined and underlined so you can identify them easily. Many comprise the common vocabulary accountants and executives use to describe the most essential dimensions of organizational environments, systems, challenges, and opportunities. Other terms are less common, but they help you draw distinctions that can mean the difference between a smart decision and a dumb one.
A solid grasp of terminology helps you and any organization you work for (or own). Here is why:

1. **Accounting is often said to be the language of business.** You can’t succeed if you don’t understand what your colleagues are telling you, or if you can’t express your own views clearly and concisely.

2. **A large vocabulary helps us see the world more clearly.** Seeing a problem is a first step toward solving it, and naming something is often the first step in seeing it. A large vocabulary gives you the eyes to see distinctions between situations that would otherwise seem similar, and to recognize similarities between situations that seem different. You can test this effect by talking with your friends about their hobbies. Avid skiers have many words to describe different types of snow ([link](#)), which help them choose whether and how aggressively to ski given current conditions. Football players have dozens of words for offensive and defensive formations to help them recognize challenges and exploit opportunities ([link](#)). Wine connoisseurs have an embarrassing wealth of descriptive terms, which help them to distinguish one wine from another by tasting them, and to assess how the wine is likely to change in quality (and value!) over the years ([link](#)). The terms in this course help you describe the most important distinctions that arise in discussions of performance and appropriate organizational structure.

3. **A large vocabulary helps us think more efficiently.** Research shows that people can keep around five to seven items in short-term memory, which limits our ability to express and evaluate complicated thoughts. But what is an “item”? People who never fly would have trouble remembering the sequence of letters DFWATLMSPLGA, because they would have to memorize twelve letters. But a frequent flyer would see this as a list of four airports (DFW is Dallas-Fort Worth, ATL is Atlanta, MSP is Minneapolis, and LGA is LaGuardia). Converting letters into airport codes is called “chunking,” a strategy that helps you remember the most important factors in your decisions and frees your brain to think about the items you actually care about: airports, not letters.

4. **Your vocabulary is one of the most visible markers of your knowledge and intelligence.** Business professionals listen, talk, and write for a living, and people judge you by the terms you know. They will judge you harshly if you don’t know a basic term, and they will judge you favorably if you can use a term to articulate a key difference between two seemingly similar situations, or a key similarity between two seemingly different situations.

**Why Are Exams “Closed Book”?**

Many students argue that they don’t need to know their material cold because they can look it up online. Even in these days of smartphones and Wikipedia, mastering the essentials without needing to refer to other material is still very useful. Here’s why:

- **You won’t always have access to your notes or Web resources—at least not without looking foolish.** Imagine that your boss asks you whether product margins are understated because the firm uses absorption accounting, and whether the firm’s accounting system is leading you into a
death spiral. Do you really want to whip out your iPhone to Google the terms? That would clearly reveal that you don’t actually know the term, and your boss will judge you accordingly.

- **Active vocabulary is more valuable than passive vocabulary.** A word is part of your active vocabulary if you can use it when it is relevant and appropriate. A word is part of your passive vocabulary if you understand when someone uses it, even if you wouldn’t have been able to think of the term yourself. Only terms in your active vocabulary help you see the world more clearly, think more clearly, and signal your expertise and intelligence. Note also that the Internet isn’t much help in maintaining your active vocabulary. If someone uses an unfamiliar term, you can always look it up. It is far harder to search for the term to use to clarify matters for yourself or for others.

- **I can’t teach if you don’t know what I am saying.** This course is cumulative—terms we learn early in the course are essential for describing the problems and solutions we cover later on. If you have to look up every term I use every time I use it in class, you won’t be able to keep up with the conversation. If I have to define every term every time I use it, we won’t get very far. So I am counting on you to become familiar with the terms when we cover them (not the night before the exam), and so are your classmates.

**Why Are Some Definitions Different from the Ones I Know?**

The vocabulary we use in this course has been developed over several centuries across many different countries and settings. As a result, sometimes a single term has many definitions (depending on whom you ask), and two terms may be defined in ways that muddy the distinction between them. In defining terms in these essays, I stick as closely as possible to their common, traditional meanings, but I modify them as needed to make them precise enough, and distinct enough from related terms, to avoid ambiguity and confusion.
Introduction

This course introduces the most important bureaucratic devices used in modern organizations: managerial reporting systems. We use these systems to track financial and nonfinancial performance in all forms of organizations. We recommend this course for students who plan to take a job at a for-profit, not-for-profit, or governmental organization of any size, or to offer one.

Three themes run throughout the course:

**Theme 1: Match Reporting Systems to Organizational Needs.** Unlike external financial reporting systems, organizations have great latitude to choose the types of systems that help them best address the challenges they face in providing goods and services, interacting with suppliers and customers, and managing intra-organizational conflict. This course introduces a wide variety of system design choices and shows how to match these choices to an even wider variety of organizational challenges.

**Theme 2: Managerial Reporting Systems Mitigate—and Are Compromised by—Intra-Organizational Conflict.** Organizations are rife with internal conflicts—between superiors and subordinates, and between peers who compete for limited resources. Good reporting systems mitigate the cost of intra-organizational conflicts, but almost all can be used as weapons in those conflicts, and as a result can become compromised. This course familiarizes students with the causes and natures of those conflicts, and with how reporting systems can address them—and be compromised by them.

**Theme 3: No System Is Perfect.** Like any bureaucratic system, management reporting systems are too crudely fashioned to provide perfect responses to every challenge an organization might face. Further imperfections arise when a system is poorly matched to its environment (theme 1), and when systems are compromised by conflict (theme 2). This course helps students recognize system imperfections, identify solutions when they exist, and accommodate imperfections when they don’t.

Underlying these three themes is a single unifying principle:

**Unifying Principle: Yes, There IS a Free Lunch.** Economists find it useful to assume that individuals exploit every possible opportunity to extract value from their circumstances. This assumption, which allows them to explain why people and organizations do what they do,
implies that there are no other opportunities for the taking—there is no “free lunch.” While the assumption is useful for economists, here is a more useful perspective for people working in business organizations (whether for-profit, not-for-profit, or governmental): of course there is a free lunch. After all, economists assume there is no free lunch only because they count on us to be sniffing out free lunches wherever we can find them.

The remainder of this essay elaborates on each of these themes.

**Theme 1: Match Reporting Systems to Organizational Needs**

The star-nosed mole is an astonishing creature, with a bright pink fleshy nose dominating its face. The picture above is from National Geographic [link], and the following description comes from Scienceblogs.com [link]. The mole’s nose has 22 fingerlike tendrils, each extremely sensitive, and able to grasp and hold as well as smell and touch.

The star-nosed mole has a unique nose because it lives in unique circumstances. The star-nosed mole lives in bogs and swamps, in a mix of mud and water. It survives by seeking out and eating small insects as rapidly as it can. It puffs air out of its nostrils and inhales immediately to sniff out its prey, making it the only mammal known to smell under water. The nose is connected so directly to its brain that it can decide within 25 thousandths of a second whether something is food or not, and grasp the food with its tendrils and place it in its mouth in not much longer. With 22 tendrils acting independently, it can sniff out and eat eight bugs every two seconds.

Organizations are complex organisms struggling for survival and success, each in its own special circumstances. Their managerial reporting systems serve as their noses, helping them sniff out opportunities and avoid challenges. Some organizations operate in highly competitive environments with razor-thin margins; their only advantage may be their ability to offer customers slightly better quality and slightly lower cost. These businesses need noses that will sniff out any opportunity to improve quality and reduce cost. Other organizations operate with little direct competition, because they customize their product or service for every client; these organizations need noses that will smell the difference between a profitable engagement and a money-loser.
Three Recurring Themes and a Unifying Principle

Theme 2: Managerial Reporting Systems Mitigate—and Are Compromised By—Intra-Organizational Conflict.

According to traditional economics, the goal of any organization is to create value greater than its component parts, while the goal of any individual is to maximize his or her individual utility. Organizations that employ more than one person invariably face some form of intra-organizational conflict, as people seek to improve their personal situations at the expense of others in the organization, and of the organization itself.

One particularly common form of conflict arises between subordinates and the superiors who wish to direct them. Superiors want subordinates to apply their effort, information, and expertise to a problem. Subordinates want to exert little effort as needed to receive as much benefit as possible from compensation and the perks (or “perks”) of their positions. Perks might include visibility, information, or connections that increase their career prospects; access to resources (a company jet, pleasant employees, and a corner office with big windows) that make their employment more enjoyable; and control over organizational decisions that determine future compensation and perquisites, along with status and other psychological benefits that most ambitious people value. Conflicts also arise among peers who are in direct competition for these limited resources.

Effective managerial reporting systems anticipate these conflicts and limit how much they hinder organizational performance. One common approach to reducing conflict is to reduce people’s ability to engage in rent-seeking behaviors. Just as landlords derive rental income from the property they own, people within organizations can derive “rental income” from the assets they control: information and the right to make decisions within the organization. Reporting systems can reduce informational rents by distributing information widely throughout the organization, and can reduce decision rents by limiting people’s unilateral powers.

People with good knowledge of how managerial systems work may increase their ability to extract rents, by influencing the design of the managerial reporting system itself to increase their informational
advantage and decision-making power. This is a particularly insidious form of rent-seeking, because few people have the expertise to understand what is happening and counteract it—to their own detriment and that of the organization.

**Theme 3: No System Is Perfect**

Good managerial reporting systems are very selective in the data they collect and report; otherwise, they would overload people with information and administrative tasks. They issue reports that simplify the world, just as a flat two-dimensional map simplifies the round three-dimensional nature of the Earth. To represent the more important parts of a map more accurately, mapmakers use a variety of “projections” that distort or delete parts that are less important. Like mapmakers, people who design reporting systems must choose which parts of their world they will misrepresent. No system is perfect, which means that designing the right reporting system for an organization is as much about deciding what it will do poorly as what it will do well.

More generally, management reporting systems are imperfect because they are bureaucratic systems: they apply standard rules every time a set of officially sanctioned conditions apply. For example, a bureaucracy might require that every expenditure of more than $5,000 be justified by a purchase request signed by the purchaser and an immediate supervisor, and that every employee receive a performance evaluation by the fifteenth day of the first quarter of every fiscal year. The systematic nature of bureaucratic systems generates imperfection for many reasons, beyond their oversimplification of the world.

- The one-size-fits-all nature of bureaucratic systems guarantees that they won’t be the best solution to every individual problem. There will surely be cases in which a bureaucracy delays an urgently needed purchase because of its paperwork requirements, or requires an unnecessary performance evaluation for an employee who has already decided to leave the firm. Nonetheless, such a system is appropriate to the broad set of problems facing a large organization when the benefit of having all purchases justified and all employees evaluated outweighs the cost of occasional inconveniences.
Three Recurring Themes and a Unifying Principle

- Bureaucracies are made even less perfect by the fact that they almost always include legacy systems, systems developed using outdated technology to solve problems that may no longer be relevant. Legacy systems make it hard for different parts of an organization to communicate, and may report information that is no longer helpful to decision-makers. Nonetheless, legacy systems are often appropriate because the benefit of using a familiar and reliable system outweighs the cost of occasional problems. Accounting itself is a legacy system, based on technology that predates the written word; it was originally designed to address the simplest problems of record-keeping. Accounting hasn’t changed much in the last 500 years; even though it hasn’t kept up with the enormous changes in business challenges and opportunities, it is good enough, given that major revisions would be difficult to implement.

- We often design bureaucratic systems to address more than one problem at a time, and therefore can’t serve either purpose perfectly. For example, we design most budgeting systems to help the production departments of a firm supply as much output as the marketing department can sell (a coordination-facilitating function), while also providing targets that managers must achieve to receive bonuses (a decision-influencing function). Unfortunately, no system can accomplish both functions perfectly, because managers who know they are going to be evaluated relative to the budget have an incentive to distort their projections of what they can produce or sell. Again, such systems are imperfect, but are appropriate when the benefits of a unified budgeting system outweigh the costs of distorted projections.

- The people who design bureaucratic systems may not have the best interests of the organization in mind. As discussed above, they may instead be engaged in rent-seeking, using their rights over system design to command additional resources.

A Unifying Principle: Yes, There IS Such a Thing as a Free Lunch!

Like so much business writing, these essays are full of lists: three themes, four functions of reporting systems, three goals of costing systems, and so on. Such lists help people remember key points, but they are admittedly somewhat arbitrary. (Why not five or two functions of reporting systems, rather than four?) But underlying all of these lists, particularly the three themes just presented, is a single principle
that unifies all of the observations and recommendations in these essays: yes, there IS such a thing as a
free lunch.

Even if you know only a little economics, this probably sounds like heresy. It directly contradicts a
widelyadmired expression popularized by Nobel-winning economist Milton Friedman: “There ain’t no
such thing as a free lunch.” (It is so widely known that commenters on many economics blogs simply use
the acronym TANSTAAFL.) The expression succinctly captures one of the most fundamental assumptions
in economics: each individual exploits every opportunity to extract value from his or her environment.
As a result, there are no easy opportunities to extract further value—the individuals in the economy
have pushed up against their boundaries, and can no longer get something for nothing. Thus, if it
appears that someone is giving away a lunch for free, economists assume that they are getting
something in return (such as a chance to make a sales pitch). Economists use this assumption to explain
and predict how market prices react to changes in technology, how organizations compensate their
workers, and a host of other phenomena.

The ubiquity of TANSTAAFL provides the foundation for a classic joke about economists:

| Two economists are walking down the street, and one asks the other, “Is that a twenty-
doller bill I see lying on the street?” The other responds: “Of course not. If it were,
someone would have picked it up!” And they walk by the bill, which is promptly
snatched up by the next passerby. |

The joke illustrates the paradox of the “no free lunch” philosophy: the only reason economists can
assume there is no free lunch is because every individual in the economy they are studying is already
sniffing out free lunches like a star-nosed mole.

This is a course for people who make decisions within and on behalf of organizations—those whose job
it is, according to economists, to sniff out every available free lunch and snap up every twenty-dollar bill
they see on the sidewalk. My hope is that this course will give you the eyes to see how to design and use
reporting systems that give you and your organization a more effective nose with which to sniff out the
free lunches that are, by economists’ own assumptions, all around you.

Epilogue: Evaluating Imperfect Systems

Sometimes a free lunch is there for the taking because an organization’s reporting system isn’t perfect.
But since no system is perfect (theme 3), how can you tell whether a proposed change to the system is
an opportunity to create value, rather than a mistake that trades off one imperfection for another that is
even worse?

Here are two useful guidelines to help you avoid making an imperfect system worse, on the one hand,
and accepting an unnecessarily flawed system, on the other.

Remember the Legacy of Chesterton’s Fence
Three Recurring Themes and a Unifying Principle

Lord G. K. Chesterton, a prolific writer of fiction, philosophy, and policy in the early twentieth century, recommended great caution in changing longstanding institutions, from religious traditions to government regulations. His advice applies equally well to organizational bureaucracies (emphasis added):

In the matter of reforming things, as distinct from deforming them, there is one plain and simple principle; a principle which will probably be called a paradox. There exists in such a case a certain institution or law; let us say, for the sake of simplicity, a fence or gate erected across a road. The more modern type of reformer goes gaily up to it and says, "I don't see the use of this; let us clear it away." To which the more intelligent type of reformer will do well to answer: "If you don't see the use of it, I certainly won't let you clear it away. Go away and think. Then, when you can come back and tell me that you do see the use of it, I may allow you to destroy it."

This paradox rests on the most elementary common sense. The gate or fence did not grow there. It was not set up by somnambulists who built it in their sleep. It is highly improbable that it was put there by escaped lunatics who were for some reason loose in the street. Some person had some reason for thinking it would be a good thing for somebody. And until we know what the reason was, we really cannot judge whether the reason was reasonable. It is extremely probable that we have overlooked some whole aspect of the question, if something set up by human beings like ourselves seems to be entirely meaningless and mysterious. There are reformers who get over this difficulty by assuming that all their fathers were fools; but if that be so, we can only say that folly appears to be a hereditary disease. But the truth is that nobody has any business to destroy a social institution until he has really seen it as an historical institution. If he knows how it arose, and what purposes it was supposed to serve, he may really be able to say that they were bad purposes, that they have since become bad purposes, or that they are purposes which are no longer served. But if he simply stares at the thing as a senseless monstrosity that has somehow sprung up in his path, it is he and not the traditionalist who is suffering from an illusion. (Link)

Chesterton’s Fence is shorthand reminding us that legacy systems were created for a reason. Before criticizing a bureaucratic system for its imperfections, we need to make sure we understand why it was initially created and why the imperfections arose. We can begin to reform the system only after we can identify the causes of the imperfection and ensure that we aren’t causing greater problems with our changes than the system forces us to endure.

Avoid Blindly Accepting the “Just World” Hypothesis

While Chesterton’s Fence cautions us to think carefully before we change systems, research on the Just World Hypothesis cautions us not to accept imperfect systems by simply assuming that they must be appropriate because they are there. After all, there IS a free lunch, if we can sniff it out.
Psychologist Melvin Lerner coined the term “Just World hypothesis” to describe a surprising set of results from his experiments: people who don’t have control over circumstances quickly learn to view those circumstances as appropriate and justified, even when they are clearly not. In one of Lerner’s experiments, subjects observed a training session in which trainees were given electric shocks each time they answered a question incorrectly. Some of the subjects were given the ability to intervene and stop the shocks. Subjects who were given the ability to intervene almost always did so, and expressed strong beliefs that the shocking was abhorrent and inappropriate. However, those who did not have the ability to intervene, and could only watch helplessly as the trainees suffered, convinced themselves that the shocks were appropriate and that the trainees “deserved” the punishment for their failure to learn rapidly enough. (You can find an interesting discussion at this link.)

Just World effects are surprisingly common. In the laboratory they cause subjects to blame people for bad outcomes that are obviously just bad luck; in the real world, Just World biases cause people to believe the poor deserve their lot and that the sick are responsible for their illnesses. Such reactions may reflect natural psychological defenses that allow humans to make sense of imperfect and unpredictable environments, and follow the tenets of Reinhold Niebuhr’s serenity prayer: “Grant me the serenity to accept the things I cannot change, courage to change the things I can, and the wisdom to know the difference.” In assessing bureaucracy, wisdom entails acknowledging the legacy of Chesterton’s fence before tearing it down, while not assuming that bureaucratic systems are right and just, simply because they seem unavoidable.
Types and Functions of Management Reporting Systems

From Financial Reporting to Managerial Reporting

Most students of managerial reporting have already studied reporting systems that generate financial statements for external parties, such as stockholders, creditors and tax authorities. While managerial reporting systems provide a source of data for external reports, their immediate purpose is to report to parties inside the firm (managers). This difference in reporting audiences has two implications:

Managerial Reporting Systems Are Largely Unregulated. External financial reporting typically must be prepared in accordance with financial reporting standards issued by the Financial Accounting Standards Board (FASB) and/or the International Accounting Standards Board (IASB). External users need to know that statements represent what they purport to represent (true and fair representation), and that similar companies represent their financial states in similar ways (comparability). Neither of these concerns arises with internal reporting. Because firms create internal reports for their own use, they have no incentive to deviate from true and fair representation. Because the user is rarely attempting to compare their performance to other similar firms, there is no need for the standards to be comparable. For this reason, there is no GAAP (Generally Accepted Accounting Principles) for managerial accounting.

Instead of financial reporting standards, management accountants are faced with a set of best practices: ways to report financial information in the most useful ways. Because best practices depend on myriad factors, including the nature of production, competitive position and organizational structure, management accounting practices are extremely varied.

Be Forewarned: Many people (and even textbooks) overstate the freedom of firms to choose their management reporting systems and the separation of internal reports from externally reported tax and GAAP income. Management reporting systems are not entirely unregulated. Government contractors need to report production costs according to guidelines set by the U.S. Federal Government, and internal control systems for many firms must be adequate to prevent financial reporting fraud and corrupt practices. Nor are internal reports completely separate from reports of taxable and GAAP income. Both the IRS and GAAP base many components of income on what a firm reports internally.

Management Reporting Systems Provide a Wide Variety of Tailored and Focused Reports. A lack of regulation is not the only factor driving variation in internal reporting practices. Variation
also arises because reports must be tailored to help managers with their particular decisions. Even a large multinational corporation like Microsoft or BP issues only a single set of financial statements combining the financial data of all of their subsidiaries over the course of a quarter or a year. Such reports are useful to investors who are trying to assess the value of an entire company, but do little to help individual managers improve financial performance. Effective management accounting systems generate reports much narrower in scope, often describing the financial performance of a single division, a single plant within that division, or a single product within that plant, over the course of a week or month. Even at the corporate level, effective reports are tailored to reflect the firm’s business strategy and highlight issues that have been of recent management interest (such as the financial cost of defective output).

Types of Management Reporting Systems

Organizations of even modest size are likely to have fairly complex managerial reporting systems that are composed of some or all of the following (sub)systems:

- **Accounting systems** that report financial information using double-entry bookkeeping.
- **Budgeting systems** that formalize an organization’s financial and operational plans.
- **Performance reporting systems** that evaluate and improve decision-making, often by comparing results from the accounting system to the budget created by the budgeting system.
- **Incentive systems** that link management reports (especially performance reports) to compensation
- **Control systems** that ensure that employees (and suppliers and customers) comply with organizational policies, and that data from the other systems are accurate.

Organizations can choose whether or not to have each of these subsystems; if they choose to have one, they then confront a wide array of choices on how to design it.

Functions of Management Reporting Systems

The goal of a management reporting system is to improve the performance of the organization and the people working in it. Most of the beneficial effects of the systems we study come through the operation of four distinct functions:

- The **attention-directing** function: inducing people to pay attention to, and talk about, one issue instead of another.
- The **decision-facilitating** function: helping people make better decisions, primarily through presenting relevant information in a format that is easy to understand. Note that decisions can be facilitated only after a manager has chosen to devote attention to making the decision.
- The **decision-influencing** function: changing the behavior of a person who is attempting to garner rewards or avoid punishment. Explicit incentives include pay-for-performance contracts; implicit incentives include possibilities for promotion, raises and dismissal based on subjective evaluations.
Types and Functions of Management Reporting Systems

- The coordination-facilitating function: helping people to coordinate their decisions, often across units.

A single element of a system often performs multiple roles. For example, a company might create a reporting system that automatically sends an email to a production manager whenever data suggests she should order additional raw materials, along with information suggesting the amount to order. This email directs the manager’s attention (by making her think about ordering inventory), facilitates her ordering decision (by providing information about suggested order quantities), influences her decision (because she will be compensated based on her ability to maintain inventory levels) and helps coordinate the production manager with her colleagues (because the email might have been generated because the sales force projected an ability to sell additional output). However, it is often useful to view these four functions as distinct, because a report may be more effective at some functions than others. For example, the email could be very effective at directing attention, but suggested order quantities might be too inaccurate to be helpful in making the ordering decision. Also, the different functions of a system may be in conflict. For example, many budgeting processes are designed to help organizations plan for the future (a coordination-facilitating function), and to specify standards against which performance can be evaluated (a decision-influencing function). These two goals are in conflict, because people won’t give us the accurate projections we need for effective coordination if they know we are going to use those projections to determine how well they need to perform in order to get a bonus. Being able to identify these two distinct functions of a single process helps us anticipate this problem and propose solutions.
Scientific Management and Taylorism

Introduction

Performance reporting systems are frequently lauded and widely adopted, but they have roots in a far more controversial practice from the early twentieth century: scientific management (sometimes called "Taylorism" in honor of its earliest proponent). This essay makes the following points:

1. Scientific management became popular because it rode a wave of scientific optimism at a time when accounting systems were available to support the data and reporting needs of the scientific method. However, it has always been controversial because it rests on two social fault lines: the conflict over the benefits of science, and the conflict between labor and management. Early proponents aggravated the controversy by signaling their allegiance to management and denigrating workers; as it has evolved from a system imposed on blue-collar workers to one imposed even on high-level managers, scientific management has become even more of a tainted term.

2. Scientific management drew its inspiration from the physical sciences. Because it largely neglects the social nature of organizational settings, scientific management leaves workers unmotivated and unsatisfied (the Human Relations Critique), it too often places decision-making rights in the hands of management and outside experts (the Organizational Critique), and it places too much emphasis on measures that are easily “gamed” (the Strategic Critique).

3. The controversies and shortcomings of scientific management and Taylorism have caused most business leaders to avoid these terms—but the practices have never been more popular. Savvy managers avoid controversy, draw from the social sciences, and recognize that scientific management will always be difficult to implement.
A Social History of Scientific Management

What Is Scientific Management?

Scientific management is the practice of applying scientific principles to production and organizational processes. Scientific management is often called Taylorism, in honor of its founder, Frederick Winslow Taylor. As a foreman at an American steel company, Taylor began to apply scientific methods to factory operations. He conducted “time-and-motion studies,” measuring exactly how much time it took workers to complete required tasks, and then holding each similar worker to the standard that repeated measurement indicated would be achievable. Taylor publicized his efforts in *The Principles of Scientific Management* (link), published in 1911, which he wrote for three reasons:

To point out, through a series of simple illustrations, the great loss which the whole country is suffering through inefficiency in almost all of our daily acts.

To try to convince the reader that the remedy for this inefficiency lies in systematic management, rather than in searching for some unusual or extraordinary man.

To prove that the best management is a true science, resting upon clearly defined laws, rules, and principles, as a foundation.

And further to show that the fundamental principles of scientific management are applicable to all kinds of human activities, from our simplest individual acts to the work of our great corporations, which call for the most elaborate cooperation. And, briefly, through a series of illustrations, to convince the reader that whenever these principles are correctly applied, results must follow which are truly astounding. (page 3, Gutenberg online version).

Taylor elaborated on his approach with an example describing four elements of scientific management in a bricklaying operation:

First. The development (by the management, not the workman) of the science of bricklaying, with rigid rules for each motion of every man, and the perfection and standardization of all implements and working conditions.

Second. The careful selection and subsequent training of the bricklayers into first-class men, and the elimination of all men who refuse to or are unable to adopt the best methods.
Scientific Management and Taylorism

Third. Bringing the first-class bricklayer and the science of bricklaying together, through the constant help and watchfulness of the management, and through paying each man a large daily bonus for working fast and doing what he is told to do.

Fourth. An almost equal division of the work and responsibility between the workman and the management. All day long the management work almost side by side with the men, helping, encouraging, and smoothing the way for them, while in the past they stood [to] one side, gave the men but little help, and threw on to them almost the entire responsibility as to methods, implements, speed, and harmonious cooperation.

Social Forces Contributing to the Rise and Popularity of Scientific Management

Scientific management became wildly popular among many leaders of large organizations—including government officials—during the first decades of the twentieth century. The movement’s popularity was motivated by a spirit of scientific utopianism and enabled by advances in accounting technology.

Scientific utopianism. The decades from 1870 to 1930 saw tremendous advances in mathematics and the physical sciences that provided the foundations of our modern understanding of the universe: set theory, the theory of evolution, the periodic table, and the theory of relativity and quantum mechanics all date from that era. Increasingly sophisticated engineering made tangible improvements in the human condition, such as the horseless carriage, the elevator, the bicycle, photography, motion pictures, radios, and medicine. It seemed that there was nothing science couldn’t accomplish, so managers naturally sought to apply the principles that had been so successful in the physical sciences to their own challenge: how to operate their businesses more profitably.

Accounting Systems. Scientific utopianism inspired managers to apply science to business, but sophisticated accounting systems provided the ability to handle the required data. As chemist Lord Kelvin (1824–1907) said, “If you can measure that of which you speak and can express it by a number, you know something of your subject; but if you cannot measure it, your knowledge is meager and unsatisfactory.” (Link to this and other Kelvin classics) Lord Kelvin’s quote is often cited as the inspiration for the common business adages “You get what you measure,” “What gets measured gets done,” and “In God we trust; all others must bring data.” However, measurement alone is hardly sufficient for good scientific management. We must reliably store the data collected by measurement, with systems to retrieve it and compile it into meaningful
Scientific Management and Taylorism

summary reports. This became possible during the era of scientific management, which saw the widespread adoption of the core systems taught in today’s management accounting textbooks, including job-order and process costing systems, standard costing systems, budgeting systems, and transfer pricing systems.

Scientific Management and Social Fault Lines

The practice of scientific management can claim credit for spawning many of the go-to disciplines for consultants, such as performance management and operations management. However, it has always been a controversial topic—enough so that savvy business leaders now avoid the tainted term “scientific management,” and “Taylorism” is now almost invariably used derogatorily (as in “Taylorism is antithetical to the humane, liberal education of young people” [link]).

Scientific management and Taylorism are controversial because they lie on fault lines that generate strong and emotional disagreements about two of the most important social issues of the modern world: disagreement about the desirability of scientific progress, and conflicts between labor and management.

Disputes About Scientific Progress. Not everyone shared in scientists’ joyous sense of discovery during the late nineteenth and early twentieth centuries. Advances in chemistry and engineering brought mustard gas and long-distance artillery in the “Great War” (World War I), which began a matter of months after Taylor published his book. Darwin’s comprehensive approach to biology inspired movements advocating eugenics at about the same time. Subsequent events cast even more doubt on the beneficial nature of science, including World War II’s specter of nuclear annihilation and the horrors of Nazi genocide. Many in the American business community had another reason to be concerned about applications of science: Marx himself viewed his theories as an unassailable conclusion of “the science of history,” rather than of philosophy or politics (link).

Disputes Between Labor and Management. The United States in Taylor’s era was fraught with labor tensions. The year 1911 saw the infamous Triangle Shirtwaist fire, which killed dozens of garment workers in New York City, followed by riots in Lawrence, Massachusetts, in 1912 (see this link for a labor-sympathetic perspective of the era). The United States then faced several decades of labor activism and the rise of powerful labor unions, as well as the establishment of Communist Party USA in 1919.

Political Insensitivity
Scientific Management and Taylorism

Taylor and other early proponents of scientific management might have had more success if they had shown more sensitivity to the social context in which they were advocating their theories. For example, it wouldn’t have taken much foresight for a good public relations agent to advise Taylor to reword this passage of his book comparing iron workers to gorillas—unfavorably—while still denying their ability to contribute to the science involved in their work.

This work is so crude and elementary in its nature that the writer firmly believes that it would be possible to train an intelligent gorilla so as to become a more efficient pig-iron handler than any man can be. Yet it will be shown that the science of handling pig iron is so great and amounts to so much that it is impossible for the man who is best suited to this type of work to understand the principles of this science, or even to work in accordance with these principles without the aid of a man better educated than he is. (Link, p30)

In another passage, Taylor notes that “the pig-iron handler is not an extraordinary man difficult to find, he is merely a man more or less of the type of the ox, heavy both mentally and physically.” Descriptions such as these betray Taylor’s very different views of labor and management. He did not mince words when describing workers’ ingenuity in shirking their duties, accusing them of working as little as they could get away with, and even conspiring to keep colleagues from showing just how efficiently work could be performed:

[I]nstead of using every effort to turn out the largest possible amount of work, in a majority of the cases [the typical worker] deliberately plans to do as little as he safely can to turn out far less work than he is well able to do, in many instances to do not more than one-third to one-half of a proper day's work. And in fact if he were to do his best to turn out his largest possible day's work, he would be abused by his fellow-workers for so doing, even more than if he had proved himself a "quitter" in sport. Under working, that is, deliberately working slowly so as to avoid doing a full day's work, "soldiering," as it is called in this country, "hanging it out," as it is called in England, "ca canae," as it is called in Scotland, is almost universal in industrial establishments, and prevails also to a large extent in the building trades; and the writer asserts without fear of contradiction that this constitutes the greatest evil with which the working-people of both England and America are now afflicted (see link above, page 7)

In contrast, Taylor argued that management would generously share the value created by scientific management, without explaining why management would not engage in the same type of rent-seeking that he thought came naturally to workers.

If you [as the owner of a shoe-making business] and your workman [has] become so skilful that you and he together are making two pairs of shoes in a day, while your competitor and his workman are making only one pair, it is clear that after
selling your two pairs of shoes you can pay your workman much higher wages than your competitor who produces only one pair of shoes is able to pay his man, and that there will still be enough money left over for you to have a larger profit than your competitor.

While it is true that you could pay the worker more, Taylor provides no reason why you would do so in the absence of market forces or collective bargaining. Taylor’s failure to justify his generous view of management amplified worker skepticism about scientific management. This skepticism has become a more serious public-relations challenge as scientific management has been imposed upon an ever-widening group of managers and professionals, including brand managers, banking executives, engineers, teachers, doctors—and MBA graduates.

Three Social-Science Critiques of Scientific Management

Scientific management and Taylorism were inspired by the success of the physical sciences at a time when social sciences were barely in their infancy. Subsequent advances in social science have revealed three shortcomings of Taylor’s theory.

The **Human Relations Critique** emphasizes that individuals have intrinsic desires to be valued as individuals, to affect their world in a positive way, to have the freedom to choose their actions, and to interact socially. These desires can lead people to work hard even in the absence of financial incentives, but can also lead them to resist changes that thwart their desires, even if those changes would benefit them financially. Scientific management that ignores these facts about human relations risks missing opportunities to draw the best from employees, while simultaneously imposing inhumane and unsustainable working conditions.

The **Organizational Critique** emphasizes that organizations function imperfectly, especially when they involve many specialized departments, many reporting layers, and complex reporting systems. Scientific management, with its reliance on complex measurement and top-down decision-making, can easily create organizational inefficiencies whose cost exceeds the savings from more efficient production. This is sometimes called the Hayekian Critique, in honor of Friederich Hayek, who emphasized the value in granting individuals extensive decision-making power in large organizations, especially governments.

Finally, the **Strategic Critique** emphasizes that measuring human behavior raises a problem that doesn’t arise in the physical sciences. Water is unaware that someone is measuring its boiling point, and won’t change its behavior as a result. But workers are usually very aware that they are being measured, how they are being measured, and why they are being measured: to determine performance evaluations, bonuses, raises, and promotions. In response, they engage
Scientific Management and Taylorism

in a practice called “measure management”: taking actions that introduce favorable biases into performance measures.

Outlook

Scientific management can be rebranded, and revised to incorporate insights from social science, but it will still be imperfect, and still create tension between those who apply the science and the “gorillas” upon whom it is imposed. However, it will continue to be widely practiced because there is no better alternative. Paraphrasing Winston Churchill’s famous quote about democracy, one could say that “Scientific management is the worst form of management, except for all the other forms that have been tried from time to time.”
Part II: Performance Reporting
Information Overload

A political strategist I know provides consulting services to candidates with small campaign budgets who are running for local offices. His clients often have grandiose visions of the messages they want to convey to voters about their plans for taxes, spending, education, defense, and social issues. The strategist’s first challenge is to rein in his clients’ ambitions; he tells them:

_You have enough money to get voters to remember two things about you…. and one of them is your name!_

This advice applies equally inside organizations, because managers receive far more information than they can possibly incorporate into their decision-making. Good reporting systems avoid such information overload through two simple strategies. First, they make managerial reports as brief as possible, focusing only on the Key Performance Indicators (KPIs). This shields managers from less relevant information that only makes it harder to pay attention to (and remember) what really matters. Second, good reporting systems organize Key Performance Indicators into a meaningful framework, so managers can easily draw them together into a memorable story about the causes of the organization’s performance.

The Balanced Scorecard

One of the first attempts to present Key Performance Indicators in a concise and meaningful way was the Tableau de Bord (dashboard), used in France since the early 1900s. Like an automobile or airplane dashboard, the Tableau de Bord provided a wide range of indicators that allowed managers to control the most important dimensions of their businesses. And the Tableau presented the indicators visually, so managers could take in the information at a glance.

The goals of the Tableau de Bord are reflected in the first version of the Balanced Scorecard, currently one of the most popular performance reporting systems in the world. Devised by Robert Kaplan and David Norton, the original version of the Balanced Scorecard groups key aspects of performance into four categories, each identified with an audience (link):
The Balanced Scorecard and Performance Reporting

- **Financial Perspective**: “To succeed financially, how should we appear to our shareholders?”
- **Customer Perspective**: “To achieve our vision, how should we appear to our customers?”
- **Internal Processes Perspective**: “To satisfy our shareholders and customers, at what business processes must we excel?”
- **Learning and Growth Perspective**: “To achieve our vision, how will we sustain our ability to change and improve?”

One goal of defining these four categories is to ensure that political pressures don’t unbalance the set of indicators. One common problem is that the accountants who create reporting systems focus too much on financial measures and deemphasize measures of marketing, operational, or human resources success. These four categories ensure that each group has “a seat at the table” when top executives or directors meet to discuss strategic objectives and performance.

A complete Balanced Scorecard also ensures that each objective is linked to a measure, a target, and an initiative:

- The **objective** is the dimension of performance the managers wish to improve. Objectives are stated in general terms, such as “Achieve high customer satisfaction” or “Maintain high yield rates.”
- The **measure** is a source of data used to assess performance on the objective. A measure of customer satisfaction might be the responses to a customer satisfaction survey, while a measure of yield rates might be the proportion of output that is free of defects.
- The **target** is a particular level of the measure to which performance is compared. For example, a firm might set a target that 90% of all customers report being “very satisfied” with the service they received, or that defect rates are less than 1%.
- The **initiative** is a particular management strategy intended to help achieve the targeted level of performance. An initiative to hit a customer-satisfaction target might include introducing a new training program; an initiative to hit a quality target might include greater incentives for engineers to anticipate and detect problems.

**Balanced Scorecard 2.0: From “Four Food Groups” to Strategy**

One shortcoming of the early incarnation of the Balanced Scorecard was the absence of a clear link to strategy. As the name suggests, the focus was balance. Thus, one might view the early Balanced Scorecard as an analog to the old “Four Food Groups” recommendations espoused by...
the US Department of Agriculture from 1956 to 1992. The USDA encouraged Americans to each some food from each of four food groups (vegetables and fruit, milk, meat, and cereals and bread). Similarly, the Balanced Scorecard encourages managers to pay attention to some measures within each of its four categories. While this Four Food Groups Approach gave marketing, operations, and human resources a seat at the boardroom table (along with accounting and finance), it suffered from the same problem as the USDA’s dietary recommendations: it was by no means clear why the groups were formed as they were, how important they were, or how they were connected.

An influential figure in quality control, M. Edward Deming, might have had this criticism in mind when he cited performance evaluation systems as one of his “Seven Deadly Diseases of Management.”1 The underlying cause of Deming’s concern is that people often create performance reporting systems “without a method of accomplishment of the objective.” A Balanced Scorecard that includes an initiative partly addresses this concern, because the initiative is exactly a “method of accomplishment.” But Kaplan and Norton went further, recasting the four perspectives into a representation of a firm’s strategic plan. In this re-release, which I call Balanced Scorecard 2.0, each perspective is a link in a causal chain. Financial performance resulted from satisfying customers, which in turn resulted from maintaining smooth and efficient internal processes, which resulted from having a trained, informed workforce with high morale.

Kaplan and Norton then supplemented this new version of the Balanced Scorecard with a strategy map showing the causal linkages among the many objectives. Spelling out causal linkages allowed managers to identify leading measures and lagging measures. Managers would expect improvements in on-time delivery to lead (occur before) improvements in customer satisfaction, but they would lag (occur after) improvements in production scheduling and quality control. This causal model addresses Deming’s concern about the absent “method of accomplishment.” It allows the strategic plan to communicate very clearly how to accomplish the stated objectives—by working harder on the factors that cause success on the objective in question.

Cascading the Balanced Scorecard Throughout the Organization

A small company with a single department may choose to implement only a single Balanced Scorecard. However, an organization with multiple organizational units (departments, divisions,

---

1 Link. Here is the full quotation of Disease #3. “Personal review systems, or evaluation of performance, merit rating, annual review, or annual appraisal, by whatever name, for people in management, the effects of which are devastating. Management by objective, on a go, no-go basis, without a method for accomplishment of the objective, is the same thing by another name. Management by fear would still be better.”
etc.) typically needs to **cascade** the scorecard throughout the organization by following these three practices:

- **Ensure that everyone in the organization understands the top-level scorecard.** The scorecard used to guide top executives at the organization captures the organization’s overall strategy. Everyone in the company, regardless of level, should know what that strategy is.

- **Provide every unit with its own Balanced Scorecard.** Every unit should have its own goals and its own strategy for accomplishing them. Thus, the scorecard is tailored to the unit’s role in the larger organization. Most units have objectives in only some of the perspectives. For example, marketing units might primarily have customer objectives, while a production unit might have mostly internal business process objectives.

- **Integrate scorecards across units.** Almost every unit in a large organization can view other units as their customers. Almost every unit has some impact on the organization’s financial performance. Every unit’s scorecard should recognize these relationships. For example, a technology services unit might view controlling total investment in hardware as a key financial objective, and reducing response time to computer problems in other units as a customer objective, but both would be internal process objectives for the firm as a whole.

### Best Practices from Governmental Performance Measurement

The United States government has devoted considerable effort to improving performance reporting by federal agencies (such as the Federal Trade Commission, the Securities and Exchange Commission, the Department of Homeland Security, etc.). The legislative powers of Congress allow them to turn best practices into the law of the land, which they did in the Government Performance and Results Act (GPRA) of 1992, which was revised by the GPRA Modernization Act of 2010. Many aspects of GPRA reflect best practices prescribed in Balanced Scorecard 2.0. However, GPRA places particular emphasis on three issues that are relevant to any organization, not just the largest one in the world.

**Draw the Target Before You Shoot**

The subjects of performance reviews often prefer a moving target—as long as they get to move it after they shoot their arrows. Here is an excerpt of a [Senate Committee Report](#) describing the importance of clearly specifying the strategic plan and associated measures, targets, and initiatives **before** performance is to be reported:

> The key to improving performance accountability is to document the results agencies have achieved compared to the goals they have established. Therefore, H.R. 2142 [GPRA]
requires an agency to provide additional information about how the agency plans to achieve its performance goals by identifying clearly defined milestones, the agency officials responsible for ensuring each goal is achieved, and the program activities, regulations, policies, and other activities that support each goal. **Without a clear description of the strategies and resources an agency plans to use, it will be difficult for Congress to assess the likelihood of the agency’s success in achieving its intended results.** By describing the strategies to be used to achieve results and the resources to be applied to those strategies, the performance plan can help Congress understand and assess the relationship between the agency’s resources and results [emphasis added].

**Clarify Each Unit’s Contribution to the Organization’s Priority Goals**

The federal government has its own goals, which are accomplished by a number of different agencies. For example, the FBI and the SEC share responsibility for detecting and deterring illegal behavior in financial markets; nearly two dozen different agencies work to prevent terrorism and collect intelligence. GPRA requires agencies to clarify exactly how they contribute to achieving these “cross-cutting federal priority goals.” From the *Senate Committee Report on GPRA*:

H.R. 2142 stresses the importance of a federal government performance plan and enhances requirements for the plan to address cross-cutting program efforts. **Focusing broadly on government-wide outcomes should be a central and distinguishing feature of the federal government performance plan.** The bill requires that:

- the plan establish performance goals for each cross-cutting federal government priority goal
- OMB [Office of Management and Budget] identify the various agencies, organizations, program activities, regulations, tax expenditures, policies, and other activities that contribute to each federal government performance goal
- a lead government official be assigned for each federal government performance goal
- OMB establish common federal government performance indicators to measure and assess progress across agencies toward shared goals
- OMB identify government and cross-agency management challenges and plans to address such challenges. [Emphasis added]

**Make Each Unit’s Priority Goals into Goals of the Organization**

A single agency might have sole (or shared) responsibility for an objective that it is important enough to be treated as an objective of the larger organization. GPRA accomplishes this by having each agency identify “priority goals,” which the Director of OMB can then designate as a “federal goal”: 
Agency Priority Goals: H.R. 2142 requires the head of each agency to identify agency priority goals from among the agency’s performance goals. The Director of OMB would have authority to determine the total number of agency priority goals across the federal government, as well as the number of priority goals to be developed by each agency.

Service Effort and Accomplishment (SEA) Reporting

The Governmental Accounting Standards Board (GASB) has also contributed to performance reporting by developing a framework focusing on the efforts and accomplishments of a governmental unit. While it is focused on governmental reporting, the SEA Reporting Framework is extremely useful for reporting the performance of service units within for-profit organizations. Like governmental units, service units are rarely in a position to report a measure of profit; instead, they probably focus on the efforts they exerted and the nonfinancial results they attained.

GASB’s Concepts Statement No. 5: Elements of SEA Reporting identifies classes of measures useful in any organization. Here are the most important:

Service effort measures include both financial and nonfinancial inputs. From GASB’s concept statement:

*Inputs*—These are measures of the amount of financial and nonfinancial resources a government uses to provide a service, such as cost of road maintenance or the cost incurred per lane-mile on road maintenance, the number of employees used in providing a service, or the amount of equipment used.

Service accomplishment measures capture both *outputs* and *outcomes*. The distinction between the two types of measures is subtle, but important. The GASB definitions are:

*Outputs*—Measures that quantify the amount of a service provided. For example, the lane-miles of road repaired, school graduation rates, number of patients treated in the emergency room, tons of garbage collected, or number of fires extinguished.

*Outcomes*—Measures that gauge the accomplishment or results that occur at least partially because of the services provided. They provide a basis for assessing how well a service's goals and objectives are accomplished. Outcome measures indicate the quality or effectiveness of a service. For example, the condition of roads in a community, or the residents' rating of the smoothness of the roads. Other examples include, the cleanliness ratings based on routine inspections of its streets or parks. As other measures that can be used to gauge success, a fire department might track the number of fire-related deaths and injuries, or the dollar value of property lost to fire. A hospital
The Balanced Scorecard and Performance Reporting

might utilize mortality rates and the results of random patient surveys. A school district might collect information on the percentage of graduating students gainfully employed or continuing education two years after graduation.

The distinction between outputs and outcomes is important, because many service providers, whether government agencies or support departments in large corporations, find it convenient to focus on outputs, which they can easily control. For example, a human resources manager who wants to do so can probably hold many training sessions for employees (an output); the manager will find it far harder to ensure that employees actually learn the content they need (an outcome). The need to influence government employees and support departments to focus on outcomes has led to a movement advocating outcome-based contracts. These contracts specify payments based on outcomes (how much did a student learn, how satisfied are residents with their roads) rather than on outputs (how many classes were held, how many potholes were filled).

Related Information That Provides Context for Understanding and Assessing Results (Optional Reading for Students)

The GASB-SEA framework emphasizes that performance measures are not sufficient to guarantee that a report is useful. Those measures must be provided in a context that helps users understand how performance compares to what might have happened, and what might have influenced the outcomes. GASB’s list of contextual items is as follows (link):

**Comparisons**— SEA performance information is particularly useful when comparisons are included, such as those with results from previous years, entity-established targets, progress towards the achievement of goals or objectives, generally accepted norms and standards, other parts of the entity, or other comparable jurisdictions (both public and private).

**Unintended Effects**— Unintended effects of a service on the recipients, state, or community sometimes can be identified. These effects represent significant indirect consequences (positive or negative) that occur, at least partially, as a result of providing a service. These unintended effects often are difficult to identify and to relate to the actual service being provided. This occurs because of an inability to establish a definite cause-and-effect relationship between the unintended effects and the service, and because other extraneous factors also may affect the results.

**Demand for Services**— Services are provided to address the needs of certain stakeholders. These needs, when recognized and expressed by stakeholders, create a demand for a service. The level of resources committed to providing such a service may reflect what is required to satisfy that need and to achieve the desired results. The level
The Balanced Scorecard and Performance Reporting

of service provided, however, may be more or less than necessary to satisfy the need addressed or the results desired. In certain instances, information about the level of need or demand for a service may help users to understand the SEA performance achieved and whether the level of services provided meets the perceived need for a particular level of service.

Factors that Influence Results—Other than the program or service delivery itself, there may be external and internal factors that influence SEA performance. Including information about these factors may help users to understand how both external and internal influences can affect results. External factors are substantially outside the control of the entity and may include economic, social, environmental, and demographic characteristics.

Narrative Information—Narrative information provided with SEA performance measures can provide explanations of what the level of SEA performance reported by the measure means, the possible effects that the factors that influence results might have on SEA performance, and actions that have been (or are being) taken to change reported SEA performance. Explanations are particularly important when reporting comparisons with other jurisdictions or among similar components within the same jurisdiction. They are also important in conjunction with reporting unintended effects of a service.
What Is Pay for Performance?

Performance reporting systems often influence behavior because they are tied to compensation. Such pay-for-performance schemes may be either explicit or implicit. Explicit incentive compensation specifies a contractual relationship between performance measures and current compensation. For example, a firm may pay an executive $500,000 in salary plus a $100,000 bonus if the firm earns at least $5,000,000, or pay a production manager $100,000 in salary plus a $20,000 bonus if defect rates are below 2%. In contrast, implicit incentive compensation, which is far more common, links promotions, raises, and noncash honors (such as job titles, “employee of the month” recognitions, and on-the-job resources and authority) to a subjective evaluation that is informed by, but not contractually tied to, performance measures.

Functions of incentive compensation

Incentive compensation schemes accomplish four purposes: motivation, communication, screening, and risk-sharing. While the first two functions are almost always intended, the last two are often unintended and can create significant problems if managers are unaware of them.

Motivation and Communication

The most obvious functions of incentive compensation are motivation and communication: pay-for-performance plans motivate workers to act in the interest of their employers, and clarify to employees what those interests are. Most of employers’ goals can be captured with two dimensions of input:

- **Working Harder** (Effort): maximizing the total physical or mental effort workers apply to tasks on the employer’s behalf
- **Working Smarter** (Effort Allocation): optimizing the allocation of effort among possible activities that benefit the employer

Working harder and working smarter are constructs that are typically difficult to measure without very expensive or intrusive monitoring. Low-level workers may be asked to punch a time clock, but even that is a poor proxy measure of effort, and sheds little or no light on effort allocation or direction. Some employers try to obtain better measures by, for example, using Radio Frequency Identification (RFID) badges that monitor time spent at a workstation as opposed to a break room or bathroom. But workers typically object to such technology, and the political costs of such technologies are rarely worthwhile if employers care about effort allocation or direction (which RFID rarely attempts to measure). As a result,
many upper-level employees (such as departmental managers and executives) may not have access to any direct measures of the behaviors their employers truly wish to motivate.

The difficulties of measuring inputs directly lead most employers to base compensation on outcomes. For example, a shirt-maker may be paid for each shirt he produces (a piece rate). The shirt-maker’s supervisor might be paid based on the number of shirts produced during his shift, adjusted for the cost or quality of those shirts. These are both measures of outputs. A service employee may be paid a bonus for achieving a high satisfaction score on a survey of her clients (an outcome measure). At the highest level, C-suite executives (CEO, CFO, COO) may be paid according to reported profits or other financial measures (outcomes).

Incentive compensation motivates workers most effectively if it follows the controllability principle: workers should be compensated based on measures over which they exercise a high degree of control. The shirt-maker typically exercises great control over how many shirts he makes in a day. As long as he is given the necessary raw materials, equipment, and workspace, he simply needs to work harder and smarter to make more shirts. It would be poor practice to try to motivate a shirt-maker at a large corporation such as Nike by paying him in shares of the corporation’s stock. The value of that stock reflects investor perceptions of the value of the entire firm, which is heavily influenced by macroeconomic factors, upper-management decisions, industry competition, and other factors far outside of the shirt-maker’s control. The shirt-maker would reasonably conclude that his actions have little effect on the stock price, and he could shirk his duties at little cost to his compensation. Even worse, stock-based compensation exposes the shirt-maker to considerable risk. Most people prefer less risk in their compensation to more (they are risk-averse), so employers who violate the controllability principle must pay workers more than if they based pay on a measure over which the employee had more control. This extra pay is a compensating differential—an increase in pay to compensate workers for taking on additional risk. (The term “compensating differential” is more commonly applied to those who face risks to their health or safety, such as coal miners or firemen; those risks are more severe than compensation risk, but the principle is the same.)

**Screening**

Incentive compensation doesn’t simply motivate workers and communicate employer goals. It also determines which workers will want to take the job. Imagine you are offered a job as a professional basketball player. Would you take a one-year contract if you were guaranteed to be paid the average salary of a player in the NBA (about $1 million)? You very well might. But what if you were paid a piece rate based on the number of points you score and shots you block? Even if you are taller and more athletic than the average business student, you probably wouldn’t expect to earn as much as you would at your next-best alternative (even if that is flipping burgers in a fast-food restaurant). You will therefore self-select not to take the job. The “self” in “self-select” emphasizes that it is your choice not to take the job. The basketball team’s owner wouldn’t need to watch you play to know that only very good basketball players would take the job—he can count on the compensation system to screen out bad players.
If the NBA example seems a bit silly, it is primarily because an applicant’s basketball-playing ability is relatively easy to assess. But the same effect operates—and is far more important—in managerial positions where it is hard to assess applicants’ abilities until they are on the job. Most salespeople are paid a low salary with high commissions based on their sales performance. Prospective employers typically face a serious information asymmetry problem, because they know far less about a job applicant’s sales ability than the applicant knows. Heavy reliance on commissions allows the employer to let the employee decide whether to take the job. Those who know they are good at sales are far more likely to want the job, while those who don’t think they are very good will prefer a job that provides a higher guaranteed salary and relies less on commissions.

Screening also works when pay is based heavily on inputs. For example, many Wall Street and legal jobs pay a great deal in bonuses or overtime to those who average close to 80 hours a week (and those who don’t may be fired). Almost everyone prefers less work to more (they are effort-averse), but some experience higher effort aversion than others. Who will take those jobs? Those who are least effort-averse!

Incentive compensation screens most effectively for ability (or effort) when performance measures are good proxies for ability (or effort). What if the performance measures are fraught with measurement error? As discussed above, basing pay on “noisy” performance measures imposes risk on the employees, and the employers must pay a compensating differential to attract risk-averse applicants to the job. Who will find the job most attractive? Those who are least risk-averse! Thus, poor performance measures screen not for ability or effort, but for risk tolerance.

Risk Sharing

The final function of incentive compensation has nothing to do with achieving better performance—it simply provides a way for employers to make their employees bear some of the risk. Entrepreneurs and business owners tend to be less risk-averse than workers, which is why they self-selected into a risky position that pays no guaranteed salary. But they are still likely to prefer less risk to more. Even if they personally don’t mind risk at all, lack of funds may force them to pay employees less when times are bad, because they must pay their other debts to avoid bankruptcy. Incentive compensation therefore serves a risk-sharing function, even if it doesn’t motivate, communicate, or screen. Thus, an entrepreneur with a new product may pay workers according to the revenue the salespeople generate, or the firm’s profits, simply because when revenue and profits are low, there is no money to pay them!

The Effort-Based Principal-Agent Model

Since the 1970s, economists have clarified the power and shortcomings of incentive compensation by modeling the employer-employee relationship as a game between two (or sometimes more) players. The terminology used in the model provides a very efficient and precise way of clarifying the assumptions hidden in the intuitive analysis above. Like any model, it dramatically simplifies the world to highlight the key tradeoffs.
Rules of the Game

The two players in the game are the principal and the agent. The principal is the owner of a firm who hires the agent to act on her behalf. (For clarity in pronouns, I refer to the principal as a woman and the agent as a man.) The game has four stages:

1. The principal proposes a compensation contract.
2. The agent decides whether or not to take the job.
3. The agent makes decisions that generate output (assuming he takes the job).
4. The principal pays the agent according to the contract, but keeps the value created by the agent’s performance (she is the residual claimant).

For simplicity, we assume that output is a linear function of pay:

\[
\text{Pay} = \text{Base Salary} + B \times \text{Performance}
\]

The parameter B measures the incentive intensity of the contract—the sensitivity of pay to performance, or (equivalently) the slope of the pay-performance function.

In stage 3, principal-agent models assume that performance is a noisy function of the agent’s effort, effort allocation, and ability. Working harder and smarter, and having higher ability, typically generate higher performance, but the agent may experience unpredictable good or bad luck. Moreover, the principal can’t distinguish luck from the other factors that determine performance. This particular form of information asymmetry (the agent knows whether he was lucky or unlucky, but the principal doesn’t) insulates the agent from the full consequences of his actions. Such insulation is often referred to as moral hazard. The term has its roots in insurance; an insurer can’t know whether a policyholder had an accident because he was careless or unlucky, so the policyholder has less incentive to be careful than he would have if he were uninsured. Similarly, the principal can’t know whether an agent was lazy or unlucky, so the agent has a bit less incentive to work hard than if he were the residual claimant on the output his effort creates.

The model includes additional assumptions about the interests of the players. The principal wants to maximize the value of the performance, after subtracting the cost of compensation paid to the agent. The agent wants to maximize compensation, after subtracting the cost of exerting effort (he is effort-averse). Finally, the agent is more risk-averse than the principal—after all, that is why he is looking for a job that pays at least some base salary, while she is an entrepreneur!

Identifying the Optimal Contract

The principal’s main problem, then, is to choose both a base salary and a level of incentive intensity. To make these choices, the principal needs to anticipate how the agent will behave in stages 2 and 3—will he take the contract, and if so, what decisions will he make? These considerations represent constraints on the principal’s choice of contract:
Paying for Performance

- The participation constraint reflects the fact that the principal needs to pay enough in total to persuade the agent to take the job.
- The incentive-compatibility constraint reflects the fact that the agent will make decisions that are in his own interest, not the principal’s.

What are the agent’s interests? The agent has an incentive to work hard and work smart because doing so generates higher performance and higher pay. However, the agent is also effort-averse—he won’t work harder unless doing so generates sufficiently higher pay. Thus, the incentive-compatibility constraint forces the principal to ramp up incentive intensity to get the agent to generate high performance. (This is the motivation function of incentive compensation.)

However, that incentive intensity comes at a cost: the agent doesn’t have full control of his performance, so higher incentive intensity forces him to share more of the principal’s risk. (This is the risk-sharing function of incentive compensation). But the agent is risk-averse, so the participation constraint forces the principal to pay a compensating differential to get the agent to take the job; the higher the incentive intensity, the higher the compensating differential.

The costs and benefits of incentive intensity also arise through its screening effects. Higher incentive intensity attracts workers with higher ability or lower effort aversion, because more able agents expect to perform better and therefore earn more incentive compensation. However, it also attracts workers with lower risk aversion, because they will not be so concerned about the uncontrollable aspects of performance. Principals might well be concerned about hiring risk-seeking agents who might put the firm in danger of failure!

Implications for Accountants

Why should accountants care about all of this economic analysis? The accountant’s job is to help management tie compensation to performance measures that are as controllable as possible! Management accountants reduce risk in performance measures in many ways. They might devise ways to monitor inputs, rather than outputs or outcomes. They might construct more timely measures, account more completely for the effects of outside factors, define the organizational unit more narrowly, or estimate financial impacts more precisely.

To see why these practices are helpful, imagine if it were possible for the principal to measure the agent’s effort with no noise at all. Then her problem would be an easy one: she would pay no base salary, and select infinitely large incentive intensity. The agent wouldn’t mind this at all, because he wouldn’t face any risk. He would simply work as hard as he could, knowing for certain that he would get paid enough to compensate him for the cost of his effort. Because the agent faces no risk, the principal doesn’t need to pay a compensating differential, and she therefore get lots of output from the agent’s hard work at a low cost. It is only when the performance measure is noisy that the principal needs to balance motivation (a good thing) and imposing risk on the agent (a bad thing).
Similarly, risk also generates the costs and benefits of screening through incentive compensation. If there were no noise, then higher incentive intensity would screen more effectively for high ability and low effort aversion, but it wouldn’t screen for risk tolerance (because there is no risk that agents need to tolerate). It is only when performance is a noisy function of effort that the principal needs to worry about incentive intensity attracting risk-seeking agents.

**Limits to the Model**

The effort-based principal-agent model is very useful in understanding the power and shortcomings of incentive compensation, but it does have an important shortcoming: it views the power of incentive compensation as working entirely because people are effort-averse, and incentives motivate them to work harder than they otherwise would. While this approach is often useful, it doesn’t work well in three situations:

*Limit 1: When Effort Doesn’t Improve Performance*

In many tasks—especially those that require creativity—more effort doesn’t necessarily generate better performance. Effort generally drives performance for rote tasks, such as data entry or ditch-digging. But puzzle-solving and artistic creation aren’t very sensitive to effort. Paying for performance on such tasks isn’t all that likely to improve performance, and some psychologists argue it even harms performance. Either way, it certainly imposes risk on the agent, because he can’t control whether he has the inspiration needed to perform well.

*Limit 2: When Agents Care About Performance for Its Own Sake*

The effort-based model assumes that the agent doesn’t care at all about the performance he generates; he only cares about how much he gets paid. In many organizations, this simply isn’t true. Consider an employee at an adoption center who has strong views on which children are appropriate to place with which parents. The principal and the agent may well have different views on such a sensitive issue, requiring the use of a disagreement-based principal-agent model, rather than the effort-based model described above. Incentives in such models are used not to motivate effort, but to alter the agent’s performance goals.

*Limit 3: When Agents Enjoy Effort*

What if a worker enjoys his job? Salespeople enjoy talking about products they sell; graphic artists enjoy designing; engineers enjoy solving problems; professors enjoy pontificating; and managers enjoy giving orders. Such people are said to have an intrinsic motivation arising from their own psychology, to distinguish it from extrinsic motivation arising from financial compensation or other benefits received from others. Extrinsic motivation not only seems unnecessary when employees have intrinsic motivation, it can even be counterproductive, due to what psychologists call the “crowding out” effect: extrinsic incentives can destroy intrinsic incentives. For this reason, psychologists warn against (for example) paying children to practice musical instruments, read, or even help around the house. Children enjoy
learning and being productive, and paying them to engage in these tasks seems to teach them that they should view these activities as chores, not hobbies.
Part III: Forms and Shadows
Do You Get What You Measure?

One of the most widely known sayings in business is “You get what you measure.” Performance reporting systems typically embrace this view by linking every objective to a measure, along with a targeted level of performance on that measure. However, not everyone shares this perspective, least of all the “father of quality control,” W. Edward Deming. Deming popularized the Plan-Do-Check-Act (PDCA) cycle during the 1940s and 1950s. One of the most influential practitioners of quality control, Deming based PDCA on the scientific method: scientists hypothesize predictions based on their theory, conduct an experiment that provides useful data, and compare the data to their predictions. In a similar way, Deming encouraged managers to plan their management or production strategy, implement it, use data to assess whether the strategy performed as expected, and then act to modify the strategy as needed.

You might think Deming would be a big proponent of “You get what you measure,” but he saw that perspective as an obstacle to good management. In the 1980s, he published an article, “The Seven Deadly Diseases of Management” (link). The list included this:

Disease # 5. Use of visible figures only for management, with little or no consideration of figures that are unknown or unknowable.

Albert Einstein emphasized a similar point (though he probably doesn’t deserve credit for the quotation, which was on a sign in his office at Princeton):

"Not everything that counts can be counted, and not everything that can be counted counts."

Do accountants have it wrong? How should management handle objectives that might be important but unknown or unknowable?

To answer these questions, we must think very carefully about the difference between (as Einstein might have said it) what counts and what we are counting. The Classical Greek philosopher Plato addressed exactly these issues, so we begin our response to Deming with a story written thousands of years ago.
What Counts and What Can Be Counted

The Allegory of the Cave

Plato was a student of Socrates, and wrote The Republic as a treatise on the nature of justice and the education required for those who would lead society. In Chapter VII, Plato used a simple story, “The Allegory of the Cave” (link), to paint a pessimistic picture of human knowledge. Like most of his writings, the story is in the form of a dialogue between Socrates and a student. Socrates begins:

[Socrates] And now, I said, let me show in a figure how far our nature is enlightened or unenlightened: —Behold! human beings living in a underground cave, which has a mouth open towards the light and reaching all along the cave; here they have been from their childhood, and have their legs and necks chained so that they cannot move, and can only see before them, being prevented by the chains from turning round their heads. Above and behind them a fire is blazing at a distance, and between the fire and the prisoners there is a raised way; and you will see, if you look, a low wall built along the way, like the screen which marionette players have in front of them, over which they show the puppets.
What Counts and What Can Be Counted

As the illustration above (link) shows, the prisoners cannot see the forms held by the puppeteers—they are chained in their chairs and cannot even move their heads. Instead, they see only the shadows cast by those forms upon the wall before them.

This distinction between forms and shadows—between true reality and our incomplete representation of that reality—is one of the earliest expressions of skepticism. Although philosophers agree on very little, most have at least some degree of skepticism, because they doubt that the true nature of the world is as it appears to us. However, Plato (through Socrates) has little faith that prisoners in his cave will share this doubt:

[Socrates] And if they were able to converse with one another, would they not suppose that they were naming what was actually before them?
[Socrates] And suppose further that the prison had an echo which came from the other side, would they not be sure to fancy when one of the passers-by spoke that the voice which they heard came from the passing shadow?
[Glaucon] No question, he replied.
[Socrates] To them, I said, the truth would be literally nothing but the shadows of the images. [emphasis added].
[Glaucon] That is certain.

Socrates next persuades Glaucon that if a prisoner were to escape into the sunlight, it would take some time before he could see and comprehend the true forms that were casting shadows in the cave. He would struggle to convince the other prisoners that their truth was “but the shadows of the images”—especially because exposure to the bright light would make it harder, rather than easier, for the newlyenlightened traveler to see and interpret the shadows in the cave as well as those who had never left. The traveler might seem to be the ancient Greek equivalent of a wild-haired Ivy League academic with his head in the clouds, disconnected from the “real world.”

One can take many lessons from the allegory, but the key insight for those interested in performance reporting systems is this: measures are merely the shadows of objectives. The grade you receive in this course is not actually how much you learned. The percentage of goods delivered on time is not actually the quality of customer service. The income reported on your tax return is not actually the increase in your wealth. These are simplified ways of counting what really counts; present-day scientists would call them proxy measures, in accordance with the common definition of a proxy as “a person authorized to act on behalf of another.” In this case, the proxy acts on behalf of the objective that we actually care about, which is called the underlying construct.

The distinction between a proxy measure and its underlying construct may seem obvious, but it is surprisingly common for people to confuse the two. One challenge is that a proxy measure is much easier to talk about than its underlying construct. Imagine that a survey of customer satisfaction yields the result that 80% of our customers report being very happy with the service they received, up from 75% last year. We know exactly what the proxy measure is: we asked a question, and customers answered it,
and our measure is the percentage of people who answer a certain way. However, we probably have a much less clear idea of what underlying construct the measure is capturing. What is “satisfaction”? Every philosopher has a different answer. But we know what question was on the survey, and we know how people responded. And we know that in the end, our pay is determined by proxy measures, not by constructs that no one can see!

A Simple Causal Model

Many accounting professors illustrate causal models using a tool called “Libby Boxes,” named after Cornell professor Robert Libby, who popularized their use for academic research. (The traditional name is the “predictive validity framework.”) Our first example illustrates a very simple causal model: “hard work leads to learning.” In this case, “hard work” and “learning” are underlying constructs that follow a very simple cause-effect relationship—hard work is the cause and learning the effect. These two constructs are shown in the top boxes, with a horizontal arrow indicating the directional effect of causality. The bottom of each construct’s box includes its proxy measure. In this case, the proxy measure for hard work is the number of hours the student reports having studied. The proxy measure for learning is the student’s score on an exam.

The oval (sometimes a fifth box) is for omitted variables—other factors that might affect the proxy measure for the final effect. Every useful model has omitted variables, because a model’s goal is to simplify the world to help explain features of interest in settings we care about. In this case, I have omitted several variables that obviously influence exam scores, such as raw intelligence, previous exposure to the material, and stress (which impairs people’s ability to form long-term memories). But these omissions are useful in highlighting the causal relation a teacher is likely to care about most—the
link between hard work and learning. We might not be able to affect intelligence, previous exposure, or stress, but we can encourage people to work harder, and expect to see that work pay off in the form of higher grades. When constructing a causal model, keep in mind another dictum from Albert Einstein:

```
Everything should be made as simple as possible, but not simpler.
```

**Measurement Error**

*Measurement error* arises whenever a proxy is an imperfect representation of its underlying construct. There are two basic forms of measurement error, *noise* and *bias*. If we think of the underlying construct as being the “true” value of something we care about (such as the true level of quality in a product, or the true effort of a worker), we can link the construct and the proxy in a simple equation:

\[
\text{Measured Value} = \text{True Value} + \text{Noise} + \text{Bias}
\]

“Noise” refers to the random and unpredictable differences between the proxy and construct. For example, imagine that our underlying construct is employees’ effort, and our measure is simply to stop by the employees’ offices once a week to see if they seem to be working. While a hard worker is more likely to be working hard during that visit, many hard workers may be taking a break when we stop by, and lazy workers may be working at that particular moment. Whether this measured value will overstate (or understate) the true value is *unpredictable*, which is why the error is noise, and not bias. Noise is the opposite of precision—statisticians often refer to a low-noise measure as having high *precision*.

Bias is a *predictable* difference between the construct and the proxy. For example, imagine that we collected measures of employee effort by calling workers in advance to tell them we were stopping by at a certain time to assess their effort level. Workers would make sure they were working hard, and our measures would be biased upward (and might be noisy as well).

The following diagram captures this more complete understanding of the relationship between constructs and proxies:
Mediating and Moderating Variables

One way to make a causal model more complex is to add one or more mediating variables. These are underlying constructs that lie between two others. For example, the Balanced Scorecard encourages managers to think of financial performance as a direct result of marketing performance, which in turn is a direct result of operational performance. Thus, one might draw a series of constructs as in the figure below. In this model, machine uptime does not itself cause customer satisfaction. Instead, it causes on-time delivery, which in turn increases customer satisfaction. Thus, on-time delivery mediates the effect of uptime on satisfaction. Note that, to avoid clutter, each construct measure is displayed with its proxy measure in the same box, differentiated by shading.
What Counts and What Can Be Counted

Model with a Mediating Variable

One can also add a moderating variable, which is a variable that increases or reduces the effect of another cause. For example, practice with a musical instrument improves performance for everyone, but the effect is bigger for those with more talent. Thus, the effect of practice on performance is moderated by talent.

Model with a Moderating Variable
What Can’t Be Counted

Note that the figure above includes a construct for “Talent,” but no associated proxy measure. It might well be impossible to measure talent directly. However, the right causal model can still help us devise ways for people to improve, if only by helping us recognize that some people need to practice far more than others to achieve the same level of performance. In contrast, we would face far more difficulty in guiding musicians if our model had incorrect causal connections (such as omitting mediating or moderating constructs or inappropriate arrows). This observation brings us full circle back to Deming’s Disease #5: ignoring unknown or unknowable variables. To avoid this disease, we must strive to represent meaningful causal models, even if we don’t expect to have good measures for our constructs.
The Difference Between Physical Science and Social Science

Assume that a manufacturing company wants to test whether a new chemical process could reduce quality problems in its factories. They might establish the effectiveness of the new process as follows. First, they identify a proxy measure of the underlying construct “quality problems,” such as “percentage of defective units.” Next, they conduct a controlled experiment, implementing the new process in some factories while keeping the old process in other, similar factories. If the experiment shows that the percentage of defective units drops more in the first group than in the second, the company is likely to conclude that implementing the new process in other factories will reduce quality problems.

Now imagine that the same company wants to reduce quality problems in its factories through a new incentive scheme, rather than a new chemical process. As before, they identify a proxy measure of the underlying construct “quality problems,” such as “percentage of defective units.” This time, however, they use that measure as the basis for a new pay-for-performance scheme. They implement the scheme in some factories by paying more to workers who generate a lower percentage of defective units, while keeping the old incentive scheme in other, similar factories. If the experiment shows that the percentage of defective units drops more in the first group than in the second, the company is likely to conclude that implementing the system in other factories will reduce quality problems.

Unfortunately, the company must be far more skeptical about the effectiveness of the incentive system than about the effectiveness of the new chemical process. The reason lies in the difference between physical science and social science that forms the basis for the Strategic Critique of scientific management: chemicals don’t care about their measured performance, but people do. There is no reason to think that the proxy measure for quality problems is more biased under the new chemical process than under the old one. But the workers under the new incentive system have an incentive not only to improve the underlying construct “quality problems,” but to improve the proxy measure itself, even when doing so doesn’t improve the underlying construct. An improvement in the proxy measure might reflect workers’ success in introducing a desired (in this case, downward) bias in the measure, relative to the underlying construct, a behavior called measure management.
Measure Management

One of the most influential social scientists of the twentieth century, Donald T. Campbell, summarized the inherently social problem of performance measurement with an observation now known as Campbell’s Law:

"The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor."

Translating Campbell’s Law into managerial reporting terms, we might say the following:

"The more incentive intensity is placed on any performance measure, the more those whose performance is being evaluated will distort the measure from the underlying construct it is intended to capture, and the more behavior will deviate from the objectives the measure was intended to promote."

Campbell’s Law expresses a specific form of skepticism, called “cynicism.” While skepticism is the belief that we cannot trust proxy measures to give reliable information about the underlying constructs we actually care about, cynicism is the belief that proxy measures are unreliable because people intentionally manipulate them.

The Law of Measure Management

Campbell’s Law suggests that measure management is driven entirely by people’s motivation. In fact, measure management arises only when three conditions are satisfied: measurement error, motivation, and ability. Exacerbating any of those conditions (not just motivation) exacerbates the problem, as stated in the Law of Measure Management.

The Law of Measure Management

Measure management is likely to arise whenever three conditions are satisfied:

1. Measurement error. The performance measure is an imperfect proxy measure for the underlying construct it is intended to capture (e.g., a performance objective).
2. Motivation. People are aware of the performance measures by which they are being evaluated.
3. Discretion. People have the ability take actions that influence:
   - a. The raw data used to compute the performance measure, (allowing real or operational measure management), or
   - b. How raw data are transformed into the performance measure (allowing opportunistic reporting).
An increase in measurement error, motivation, or ability will tend to increase the likelihood and severity of measure management.

In the manufacturing firm described above, “percentage of defective units” is unlikely to be a perfect proxy for quality problems, especially if the firm makes a variety of products that have different levels of difficulty (and therefore different minimum achievable defect rates). If the managers being evaluated can choose the mix of different products, they can engage in real measure management (or, synonymously, operational measure management) by choosing to produce only the easy products. This will lower reported defect rates, but it doesn’t reflect actual improvement in quality processes at the plant. Alternatively, if the managers have discretion over what constitutes a “defect,” they might engage in opportunistic reporting by lowering their standards for what products meet product standards. Either form of measure management reduces reported defect rates relative to the true level of quality in the plant.

Examples

Some of the most common and widely publicized forms of measure management arise in external financial reporting, where it is usually called “earnings management” (management in which earnings is the measure managed). Countless studies document that firms engage in real operational earnings management by timing their sales and purchases to generate desired earnings results. Well-known scandals (Enron, WorldCom, Health South, etc.) result from forms of opportunistic reporting in which firms use their discretion to report outside the bounds of GAAP (for example, WorldCom calling certain expenditures “assets” rather than “expenses,” as required by GAAP). Note that both forms of earnings management still satisfy the more general definition of measure management, because they cause the reported measure “net income” to be a biased representation of the underlying construct it is intended to capture: the true change in the value of the firm due to operations.

Earnings management is a complex topic that requires detailed understanding of the mechanics of corporate transactions, GAAP, and double-entry bookkeeping. We address some examples of earnings management later in the course, and other courses in the school (Intermediate Accounting and Financial Statement Analysis) explore them in even greater detail.

To provide more accessible illustrations of the Law of Measure Management, I use examples from two settings that require less technical knowledge, and are topics of current political controversy: education and medicine.
Nichols and Berliner (2005) provide a comprehensive survey of measure management in education in their report, “The Inevitable Corruption of Indicators and Educators Through High-Stakes Testing” (link). Under recent federal laws, public schools in the U.S. are evaluated on their students’ standardized-test scores. Educational success is notoriously difficult to measure, so standardized tests measure that construct with more than enough error for measure management to be a concern. Teachers and administrators also have strong motivations to consider such behavior:

Standardized-test scores and other variables used for judging the performance of school districts have become corruptible indicators because of the high stakes attached to them. These include future employability of teachers and administrators, bonus pay for school personnel, promotion/non-promotion of a student to a higher grade, achievement/non-achievement of a high school degree, reconstitution of a school, and losses or gains in federal and state funding received by a school or school district. [Executive Summary]

Schools employees use their discretion over teaching content to engage in real measure management by “teaching to the test” and narrowing the curriculum:

One fifth-grade teacher says that “the state's rigorous testing schedule is failing students because it leaves teachers no choice but to base instruction on what's being asked on the tests.” "Tests have become the holy grail of education," said Wiener, a teacher at Ynez School in Alhambra. Because he has to focus so much on math, reading, and language arts, Wiener has had to "downplay many other subjects, especially the arts. "I think public speaking is important, but it takes up a lot of time and it's not on the test," he said. Weeks can go by without students having art and music classes, he said. [Table 6]

Administrators also use their discretion over class composition to engage in real (operational) measure management by discouraging test-taking by students who are likely to perform poorly:

In Tampa, a student who had a low GPA and failed portions of the state’s standardized exam received a letter from the school encouraging him to drop out even though he was eligible to stay, take more courses to bring up his GPA, and retake the standardized exam....

In New York, thousands of students were counseled to leave high school and to try their hand at high school equivalency programs. Students who enrolled in equivalency programs did not count as dropouts and did not have to pass the Regents’ exams necessary for a high-school diploma.” [Executive Summary]
The report includes numerous examples of opportunistic reporting, in which school employees use their ability to influence the raw data that determine test averages—students’ answers, and in a more complex case, student identification information:

Helen Lehrer, the former principal at Pacific High School in Boerum Hill, was accused of erasing and changing as many as 119 answers on 14 Regents competency tests in global studies and U.S. history given last June.

[In other schools]... “student identification number changes were submitted for students tested at [the schools], which resulted in the exclusion of those students from the accountability subset of TAAS results used to determine the 1998 accountability ratings.” In plainer English, administrators gave students who performed poorly on the test ID numbers that did not agree with previous numbers, knowing that the inconsistencies would cause the TEA to eliminate the students’ scores from ratings calculations. [Table 1]

**Medicine**

An article by Werner and Asch in the *Journal of the American Medical Association* discussed some of the unintended consequences of publicly reporting information about hospital quality. Like education, medical quality is extremely hard to measure. Werner and Asch describe the state of reporting as follows:

Two general types of health care report cards exist: those that measure outcomes and those that measure process. Reports of cardiac surgeons’ and hospitals’ risk-adjusted mortality rates following coronary artery bypass graft (CABG) surgery are examples of outcomes-based reporting. Process-based report cards, often called quality indicators, report on rates of medical interventions, such as screening tests and medication use, which are assumed to be related to outcomes.

The report cards motivate measure management because patients are likely to choose those who report strong performance over those who report weak performance; hospitals generating insufficient performance may find themselves unable to qualify for preferred status among private insurers and governmental programs.

Werner and Asch examine the history of the Coronary Artery Bypass Graft (CABG) report card in New York State. Their analysis shows that, while releasing the report card improved reported
performance, the result was at least partly due to both real measure management and opportunistic reporting:

After New York began releasing its CABG report card, CABG mortality rates in New York dropped from 3.52% in 1989 to 2.78% in 1992, a decrease of 41%. This decline in CABG-associated mortality rates was larger in New York than in other states at the same time and persisted through the 1990s. The mortality decline was hailed as a success of the CABG report card.

However, this enthusiasm has been curbed by simultaneous reports of cardiac surgeons turning away the sickest and most severely ill patients in states with CABG report cards in an effort to avoid poor outcomes and lower publicly reported ratings. Omoigui et al. noted that the number of patients transferred to the Cleveland Clinic from New York hospitals increased by 31% after the release of CABG report cards in New York, and that these transfer patients generally had higher risk profiles than patients transferred to Cleveland Clinic from other states. In Pennsylvania, which also introduced CABG report cards, 63% of cardiac surgeons admit to being reluctant to operate on high-risk patients, and 59% of cardiologists report having increased difficulty in finding a surgeon for high-risk patients with coronary artery disease since the release of report cards. New York had a similar experience after the release of report cards, reporting that 67% of cardiac surgeons refused to treat at least 1 patient in the preceding year who was perceived to be high risk.

Moreover, patients undergoing bypass surgery in Pennsylvania and New York had lower illness severity than patients in states that did not publicly release the information, particularly among surgeons rated as low quality. Furthermore, the release of New York’s CABG report card was associated with an increase in racial disparities in CABG use, suggesting that surgeons also may have responded to CABG report cards by avoiding patients perceived to be at risk for bad outcomes, such as blacks and Hispanics. Although some prior studies have noted improvements in CABG mortality rates after the release of CABG report cards, if quality report cards cause physicians to select patients based on risk profile, the quality of care and outcomes of people eligible for CABG may worsen even as mortality rates among those who receive CABG improves.

The Legality and Ethics of Measure Management

The term “measure management” refers to activities that intentionally and effectively introduce a favorable bias into a proxy measure, relative to the underlying construct it is intended to represent. Calling a behavior “measure management” does not imply that it is illegal, unethical, or even undesirable. Some practices described above clearly violate laws,
Measure Management

regulations, or employment policies (such as changing students’ test answers), whereas others are clearly legal (such as devoting less time to music education, teaching to the test, or declining to operate on a high-risk patient). One can surely debate the ethics or desirability of these practices, but the fact that a practice is appropriately called “measure management” plays a minor role in such assessments.

We discuss the ethics of measure management in detail later in the course.

Responses to Measure Management

Measure management is difficult to avoid, but the key to reducing it lies in the Law of Measure Management that describes when it is most likely to arise:

- **Reduce measurement error.** The more accurately your measures capture the constructs you care about, the less likely it is that managing to the measure will create a wedge between measure and construct.
- **Conceal measurement.** People who don’t know they are being measured can’t easily distort their operations or report opportunistically. Restaurant critics and “secret shoppers” reduce measure management by concealing their identity and making sure they are treated the same as any other customer.
- **Monitor and limit discretion.** People can’t engage in measure management if they don’t have the ability to distort real operating decisions or how performance is reported. Monitor opportunities for real measure management (such as declining high-risk patients at a hospital), and make sure managers know that such behavior will be punished. Limit reporting discretion by having information processed by people other than the ones being evaluated, or if that is impossible, make sure people know that their reporting choices are being monitored and that overly opportunistic reporting will be dealt with harshly.
Part IV: Accounting Basics
Introduction

Debits, credits, and other mechanics of double-entry bookkeeping strike fear into the hearts of many. Those essential accounting topics are easier to follow if one first understands the fundamental building blocks of accounting: the elements of financial statements that accountants use to describe wealth and how it changes from one period to the next. The elements of wealth, shown on a firm’s balance sheet, are assets, liabilities, and equity. The elements of changes in wealth are revenues, expenses, gains, losses (shown on the income statement), and contributions from and distributions to owners (shown on the statement of changes in shareholders’ equity).

Tracking Wealth

History provides a gentle and progressive introduction to the underlying constructs that accountants seek to track. People developed the earliest accounting systems in ancient Babylonia and Egypt. They attempted to do little more than list assets, which four or five millennia later the Financial Accounting Standards Board (FASB) defined as follows (this and subsequent definitions of elements are taken from the FASB’s Statement of Accounting Concepts No. 6: Elements of Financial Statements) (link).

*Assets* are probable future economic benefits obtained or controlled by a particular entity as a result of past transactions or events.

Two words in the definition are worth additional emphasis. First, the term *entity* refers to whose wealth is being tracked. Even in ancient days, an entity might refer to any number of social unit: an individual, a family, or a government. Legal advances have led to the creation of many new entities, the most important of which are the corporation, the partnership, and the trust. Each of these entities is distinct from any person, even if they are owned by a single person.

The word *controlled* takes on pointed meaning in the ancient world: one must be able to protect one’s property rights through force of law or arms. For a farmer, an asset might have been livestock that the farmer could protect himself, or could expect to be protected by a
Elements of Wealth and Income

government that defined and outlawed theft. For early governments, assets also included the ability to confiscate residents’ assets through taxation. The government’s tax claims naturally required quite a bit of record-keeping, and tax records are not only the earliest forms of accounting, but even formed the basis for writing itself. Denise Schmandt-Besserat, a French archeologist, has argued that cuneiform, the first known form of writing, was driven by the need to keep track of wealth and taxes. As she writes in her influential book *How Writing Came About* (link),

The immediate precursor of cuneiform writing was a system of tokens. These small clay objects of many shapes—cones, spheres, disks, cylinders, etc.—served as counters in the prehistoric Near East and can be traced to the Neolithic period, starting about 8000 B.C. They evolved to meet the needs of the economy, at first keeping track of the products of farming, then expanding in the urban age to keep track of goods manufactured in workshops. The development of tokens was tied to the rise of social structures, emerging with rank leadership and coming to a climax with state formation.

You’ll have to forgive accountants for their pride, but it is well deserved!

Assets arising from claims to others’ property leads naturally to a second fundamental concept: that of the liability:

**Liabilities** are probable future sacrifices of economic benefits arising from present obligations of a particular entity to transfer assets or provide services to other entities in the future as a result of past transactions or events.

With the concepts of assets and liabilities well-defined, one can then build a construct of wealth. In plain English, your current wealth is what you have (assets) minus what you owe to others (liabilities):

**Equity**, or **net assets**, is the residual interest in the assets of an entity that remains after deducting its liabilities.

This definition of wealth is also called “net assets” or “equity,” and is captured in the fundamental accounting identity:

\[
\text{Assets} = \text{Liabilities} + \text{Equity}
\]

**Tracking Changes in Wealth**

Definitions (and lists) of assets and liabilities are sufficient to describe wealth at a moment in time. For this reason, we call them **stock variables**. But advances in civilization motivated
people to consider how their net assets changed over time. As early as Renaissance times, people found it useful to distinguish between two reasons net assets might change. First, the net assets of an organization might increase because the owners invested new assets in their business or might decline because the owners took a portion of the wealth for their own purposes. Today such additional investments are typically called “contributions,” while payments are called “dividends” (for corporations) or “distributions” (for partnerships). All of these payments fall in the category of “transactions with owners.”

[Contributions]² by owners are increases in equity of a particular business enterprise resulting from transfers to it from other entities of something valuable to obtain or increase ownership interests (or equity) in it.

Distributions to owners are decreases in equity of a particular business enterprise resulting from transferring assets, rendering services, or incurring liabilities by the enterprise to owners. Distributions to owners decrease ownership interest (or equity) in an enterprise.

How else might an entity’s net assets change? By creating or losing wealth. Such changes in net assets form “comprehensive income”:

Comprehensive income is the change in equity of a business enterprise during a period from transactions and other events and circumstances from nonowner sources. It includes all changes in equity during a period except those resulting from investments by owners and distributions to owners.

Contributions, distributions and comprehensive income are called flow variables because they describe changes in equity from one moment in time to another (usually a month, quarter, or year later). The distinction between Comprehensive Income and the other flow variables was crucially important in the Middle Ages, when wealthy families would provide income to their heirs in the form of trusts. Under such arrangements, an heir would be entrusted with the assets of an ancestor, but would be permitted to extract capital only to the extent of income. In this way, one could write a will to guarantee that one’s wealth would be bequeathed from generation to generation, because no generation could extract more wealth than they created through comprehensive income.

² The FASB uses the term “Investments.”
Components of Income

Revenue vs. Gains

As the complexity of business arrangements increased, accountants found it useful to distinguish among the many different components of comprehensive income. The term “revenue” describes an increase in wealth due to the receipt of consideration from a customer in the ordinary course of business operations:

Revenues are inflows or other enhancements of assets of an entity or settlements of its liabilities (or a combination of both) from delivering or producing goods, rendering services, or other activities that constitute the entity’s ongoing major or central operations.

Other increases in wealth were termed “gains”:

Gains are increases in equity (net assets) from peripheral or incidental transactions of an entity and from all other transactions and other events and circumstances affecting the entity except those that result from revenues or investments by owners.

The distinction is important because revenue is usually recurring: the ability to persuade a customer to part with cash in exchange for an entity’s asset suggests that the business has a continuing source of cash (as long as it can keep providing the product or service customers value). However, a gain, an increase in wealth due to some other event (such as a profitable investment in land or a successful gamble) may well be a one-time event.

Note that revenue need not be the outcome of a transaction. Some businesses, such as mining, agriculture, and farming, create new assets on a recurring basis without transacting with customers. By the definitions above, the birth of a lamb would reflect revenue, because it is an ongoing activity and is therefore likely to be followed by more wealth creation. In contrast, a lamb that wandered from the wild onto your property would be a gain, since such an event is rarely repeated.

Expenses vs. Losses

A business can lose wealth for two reasons that are analogous to the two forms of income. A firm can lose wealth in the course of its ordinary operations, as it seeks to create revenue or recurring income (such as raising livestock). Such decreases in wealth are referred to as “expenses.”
Expenses are outflows or other using up of assets or incurrences of liabilities (or a combination of both) from delivering or producing goods, rendering services, or carrying out other activities that constitute the entity’s ongoing major or central operations.

Other decreases in wealth are losses, and are analogous to gains:

Losses are decreases in equity (net assets) from peripheral or incidental transactions of an entity and from all other transactions and other events and circumstances affecting the entity except those that result from expenses or distributions to owners.

Components of Assets and Expenses: Cost Accounting Edition

Total assets can be broken down into a variety of components; two of these are essential for understanding how to account for the costs of operating a business. The FASB’s Concept Statement does not provide formal definitions for these terms; therefore, all of the definitions that follow are my own.

The first asset component is inventory:

Inventory includes assets that are held for sale to customers or for conversion into an asset that is held for sale to customers.

The three key inventory accounts are Raw Materials (RM), Work-in-Process (WIP), and Finished Goods (FG). Working backward, Finished Goods includes assets immediately ready for sale to customers; Work-In-Process includes assets that have been partially processed but are not yet complete; Raw Materials are assets that have been purchased but not processed at all. Thus, RM is converted into WIP, which in turn is converted into FG.

The second key component of assets for cost accounting is Property, Plant, and Equipment (PP&E):

Property, Plant, and Equipment includes physical assets that are vital to ongoing operations, but are not held for sale to customers or converted into an asset held for sale to customers.

As the name suggests, PP&E includes items such as a factory or office building, the land on which a factory or office building is housed, and equipment (such as machines, cars, tools, computers, and the like) used in ongoing operations.

To the cost accountant, the most important distinction between components of expenses is that between Cost of Goods Sold (CGS) and Selling, General, and Administrative Expenses (SG&A):
Cost of Goods Sold is the expense incurred because the entity has parted with goods or services conveyed to customers as part of ongoing operations.

Selling, General and Administrative Expenses are all expenses incurred to maintain ongoing operations other than Cost of Goods Sold.

The primary component of CGS is the value of FG transferred to customers (which in turn is driven by the value of WIP, which includes the value of RM and processing costs). Costs that pass through FG into CGS are also called product costs, because they are directly associated with the product provided to customers. SG&A includes costs of upper management and other back-office administration, marketing, advertising, and so on. Costs that are passed into SG&A are also called period costs, because they are incurred simply to conduct operations over the reporting period.
The Romance of Accounting

Accounting elements describe wealth at a moment in time, and changes in wealth from one moment to the next. The Renaissance saw the development of double-entry bookkeeping, which ties these elements together in an elegant and useful way. While most people think of double-entry bookkeeping as a pedestrian achievement, one of the most renowned figures of the early Romantic era had a different perspective:

Double-entry bookkeeping is one of the most beautiful discoveries of the human spirit.... It came from the same spirit which produced the systems of Galileo and Newton and the subject matter of modern physics and chemistry. By the same means, it organizes perceptions into a system, and one can characterize it as the first Cosmos constructed purely on the basis of mechanistic thought.... Without too much difficulty, we can recognize in double-entry bookkeeping the idea of gravitation, or the circulation of the blood and of the conservation of matter.

Johann Wolfgang Von Goethe, 1796

Accounting may be the butt of many jokes, but it is a proud and noble profession!

Accounts in Balance

The central insight of double-entry bookkeeping is that the elements of wealth and income can be captured in accounts that are kept in balance by changing their values in tandem. The “Queens Cross,” devised by Accounting faculty at Queens University and shown in Exhibit 1, elegantly communicates the relationships between these accounts. The top of the cross houses assets on the left and liabilities and
equity on the right. The bottom of the cross houses expenses, losses, and distributions on the left and revenues, gains, and contributions on the right.

Exhibit 1 shows the values of a sample set of accounts at the beginning of an accounting period for a simple business. Note that the dollar value of the assets equals the dollar value of the liabilities plus equity, as required by the fundamental accounting identity. The top of the Queens Cross reflects the firm’s Balance Sheet at a specific point in time (in this case, 8 A.M., January 1, 2011).

<table>
<thead>
<tr>
<th>Assets (as of 8am, 1-1-2011)</th>
<th>Liabilities &amp; Equity (as of 8am, 1-1-2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash</td>
<td>0</td>
</tr>
<tr>
<td>Accounts Receivable</td>
<td>$1000</td>
</tr>
<tr>
<td>Raw Materials</td>
<td>$750</td>
</tr>
<tr>
<td>Finished Goods</td>
<td>$250</td>
</tr>
<tr>
<td>Equipment</td>
<td>$1000</td>
</tr>
<tr>
<td>Debt</td>
<td>0</td>
</tr>
<tr>
<td>Equity</td>
<td>$3,000</td>
</tr>
</tbody>
</table>

Now assume that the following events take place over the course of the day, January 1, 2011:

A. The business spends $150 of its cash to buy raw materials. This event transforms one asset into another. To keep the books in balance, we reduce the cash account and increase the raw materials account to keep the left and right sides in balance, and only accounts in the top left quadrant are affected.

B. The business sells its finished items to a customer for a promise to pay $500. This actually triggers two different events: the business gives up its finished goods inventory, and it receives a promise to pay.
   a. The sacrifice of inventory meets the definition of Cost of Goods Sold Expense, so we reduce the finished goods inventory account and increase the Cost of Goods Sold account in the lower left quadrant, keeping the left and right sides in balance.
   b. Since the promise from a customer for the delivery of goods satisfies the definition of Revenue, we increase Accounts Receivable by $500 and increase the Revenue account by $500 in the bottom right quadrant, again keeping the left and right sides in balance.
C. A machine breaks down and cannot be repaired. This generates a decline in wealth of $50, because we had recorded the machine as an asset, and it no longer provides any value. We must therefore reduce the value of the Equipment account. The decline satisfies the conceptual definition of a loss, so we debit a loss account, keeping the left and right sides in balance.

D. The business promises to pay its accountant $100 for doing such a good job tracking its wealth and income for the day. We could keep the left and right side of the Queens Cross in balance by recording a liability (Wages Payable) and reducing equity, but this would conceal the reason equity declined—the business sacrificed assets in the ordinary course of operations for administrative expenses. Since this sacrifice meets the definition of Selling, General, and Administrative expenses, we record a liability and an SG&A expense. This still keeps the left and right in balance, but adds information about the nature of the event.

The following Queens Cross shows the accounts after all of these events are recorded. We now have the makings of an income statement for the day 1-1-2001, which would show income of $100 (Revenue of $500 less expenses and losses of $400).

![Queens Cross]

But we can’t stop quite yet. Note that the balance sheet violates the fundamental accounting identity because the assets do not equal liabilities plus equity. We can fix this by zeroing out the accounts on the bottom of the cross, and balancing those changes with changes in equity. Because the expense reflects a decrease in wealth, we lower the expense to zero and lower the equity account by the same amount. Because the revenue account reflects an increase in wealth, we lower the revenue to zero and increase the equity account by the same amount.
Now that the income accounts have been zeroed out ("closed"), we can present a new balance sheet, as of 5 p.m., January 1, 2011:

<table>
<thead>
<tr>
<th>Assets (as of 5pm, 1-1-2011)</th>
<th>Liabilities &amp; Equity (as of 5pm, 1-1-2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash $850</td>
<td>Debt $100</td>
</tr>
<tr>
<td>Accounts Receivable $500</td>
<td>Equity $3,100</td>
</tr>
<tr>
<td>Raw Materials $900</td>
<td></td>
</tr>
<tr>
<td>Finished Goods $0</td>
<td></td>
</tr>
<tr>
<td>Equipment $950</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expenses &amp; Losses For the day 1-1-2011</th>
<th>Revenue &amp; Gains For the day 1-1-2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of Goods Sold 0</td>
<td>Revenue 0</td>
</tr>
<tr>
<td>SG&amp;A 0</td>
<td></td>
</tr>
<tr>
<td>Loss on Equipment 0</td>
<td></td>
</tr>
</tbody>
</table>

**Bookkeeping Operations**

The preceding introduction describes what happens to the various accounts, but it doesn’t shed much light on how bookkeepers would actually make these changes in a way that allows them to leave a paper (or electronic) trail that can be audited by others to ensure its accuracy.

A Renaissance mathematician and scientist, Friar Luca Pacioli, summarized the operations of bookkeeping in the late 1400s, and they have changed little since then. The accountant’s first task is to record all key events that have occurred in the business that are relevant to the entity’s wealth. This record would include every transaction with outsiders, along with any key business events that might alter the value of existing assets and liabilities. In the present day, we would draw these records from a computerized system tracking sales, production, and administrative activity, ideally in real time in an integrated Enterprise Resource Planning system like SAP or Oracle.

To convert these records into account balances, accountants use a very special convention regarding left and right. Reflecting bookkeeping’s Renaissance roots, we still call left and right by their Latin names: “debit” and “credit.” The rules of left and right are simple:

- To debit means to make an entry on the left; to credit means to make an entry on the right.
- Accounts on the left side of the Queens Cross (assets, expenses, losses, and distributions) increase when debited and decrease when credited; accounts on the right side of the Queens...
Cross (liabilities, revenue, gains, and contributions) increase when credited and decrease when debited.

The entries for the events are then recorded in a list of journal entries, which are still formatted largely as they were in the 1400s. Here are the entries for the four events used in the example in the essay:

**Accounting Elements:**

**Entry 1**

Raw Materials $150  
Cash $150  
*To record the purchase of raw materials for cash*

**Entry 2a**

Cost of Goods Sold $250  
Finished Goods Inventory $250  
*To record expenses from the sale of Finished Goods Inventory*

**Entry 2b**

Accounts Receivable $500  
Revenue $150  
*To record revenue from the sale of Finished Goods Inventory*

**Entry 3**

Loss on Equipment $150  
Property, Plant, and Equipment $150  
*To record the breakdown of equipment*

**Entry 4**

Selling, General, and Administrative Expenses $100  
Wages Payable $100  
*To record the cost of accounting services*

Next, the accountant must record the entries into a ledger, which lists the initial balances and changes to every account. Many bookkeepers use the “T-account” to capture such information. In a T-account, the account name is shown at the top, and the initial balance is shown unnumbered at the top of the left or right, depending on whether the initial balance is a debit or credit. Every change is also recorded as a debit or credit, along with a letter or number that refers to the associated journal entry. The last page of this document presents the entries in T-account form. Note that we omit the original balance because it is irrelevant to the purpose of an exercise (which is to emphasize the entries that record the transactions, not the final balances).
Three Reasons to Make Entries

There are three reasons to make a journal entry: to reflect a transaction (a transaction entry), to adjust an account balance to reflect a change in circumstances other than a transaction (an adjusting entry), and to close out an account that needs to have a zero balance before the next reporting period begins (a closing entry). We have seen examples of all three types of reasons (and associated types of entries) in this simple example.

- The entries to record the purchase of raw materials, the sale of finished goods, and the payment to the accountant were all transaction entries. Note that transactions need not be with a party outside the firm—a transfer of goods from one department to another (e.g., Raw Materials to Work in Process) is also a transaction. Best practice is to record a transaction as soon as it occurs, because all of the necessary information is typically available to the bookkeeper and delay produces no benefit.

- The entry to record the breakdown of the machine is an adjusting entry. An event occurred that reduced wealth, but it was not a transaction. Best practice is to record adjusting entries at the end of the accounting period, because the true wealth effect of the event is likely to become clearer as time passes.

- The entry to zero-out the income statement balances and make corresponding entries to equity are closing entries. These entries must be made at the very last second, because all transactions and adjustments must first be determined before the accounts can be set to zero.
Double-Entry Bookkeeping, Simplified

Assets
- Cash: A 150
- Raw Materials: A 150
- Equipment: C 50

Liabilities/OE
- Accounts Receivable: Bb 500
- Wages Payable: D 100

Equity
- Finished Goods Inv.: Ba 250

Expenses/Losses
- Cost of Goods Sold: Ba 250
- Selling, General & Admin: D 100
- Loss on Equipment: C 50

Revenues/Gains
- Revenue: Bb 500
Belief, Skepticism and Postmoderism in Accounting

Accounting Forms and Shadows

The three quotations in the joke above reflect different perspectives on how reported outcomes relate to reality.

- The first umpire is a true believer in his abilities to make the right call. He believes that he can perfectly detect the location of a three-inch baseball during the fraction of the second that it passes over home plate, even though it curves unpredictably while moving faster than a car on a freeway. Players and their fans are likely to disagree!

- The second umpire is a classic skeptic. (Perhaps he has read What Counts and What is Counted.) This umpire recognizes that he does the best he can to shout “strike” if the ball appears to have crossed the plate within the strike zone, and shout “ball” otherwise. He would probably admit to his calls being noisy shadows of the underlying form of each pitch, even if he wouldn’t admit that his calls are tainted by biases of the type shown in this link.

- The third umpire is what philosophers call a postmodernist, because he denies that that there is any underlying reality that stands independent of how people report it. A pitch is neither a ball nor a strike in the absence of this umpire’s report on the matter. He is probably modeled on Bill Klem, umpire for Major League Baseball from 1905 to 1941; see this link for more. Klem’s famous claim that “it ain’t nothin’ till I call it” was a strong assertion of his powers to dictate outcomes, but didn’t shed much light on how he came to his decisions.

The three umpires reflect three common views about accounting reports. Imagine that the athletic apparel giant Nike spends millions of dollars to bolster their “Just Do It!” marketing campaign. Just as
Belief, Skepticism and Postmoderism in Accounting

the umpire must call a pitch a ball or a strike, an accountant must decide whether to debit an asset or an expense. The accountant might echo the first umpire by saying “This costs of this campaign are either an asset or an expense, and I can perfectly determine which.” The accountant might echo the second umpire by saying “These costs are either an asset or an expense, and I’ll do the best I can to determine which.” Or the accountant might take a postmodern approach, and claim that by his report, he renders the cost either an asset or an expense; before he decides, it “ain’t nothin’”.

One challenge in teaching Managerial Accounting is that students often leave their course in Introductory Financial Accounting as either believers or postmodernists. In the U.S., such courses focus primarily on how the FASB has decided to account for each common transaction. The FASB has decided, for example, that cash spent on marketing or research and development should be expensed immediately. They have also decided that firms should recognize losses if the market value of an asset has declined below book value during the period, but they should not recognize a gain if the market value has increased above book value (unless it is a special class of financial assets). After many weeks of such training, students stop thinking about the difference between accounting forms and shadows, which doesn’t help them earn a higher grade. Slowly but surely, the come to believe unquestioningly that the FASB’s rules infallibly capture accounting reality, or (if they are more cynical) accept that the notions like asset and expense have no meaning beyond how the FASB uses them.

Successful students of managerial accounting must quickly learn to be skeptics. Just as there are no atheists in foxholes, there are no believers or postmodernists in an organization that is battling over how to report financial performance—and organizations face these battles regularly. Financial accountants can outsource their battles to the FASB (which, by the way, is a highly political organization, and battles rage on for decades). But inside the organization, managers and accountants must wage their own battles.

For example, imagine you are overseeing a marketing campaign for the launch of a game-changing product. The accountants at your firm immediately expense all of your spending. While they must do this for external reporting (because the FASB demands it), they have the freedom to report that spending as an asset for the internal reports on which your performance is evaluated. Is it right that your activities are reported as a loss of wealth for the firm? Won’t your campaign pave the way for high sales in future years? Or imagine that you are responsible for buying property that will serve as locations for your firm’s chain of retail stores. Due to your excellent judgment, the value of the property you bought has increased substantially in value. However, the firm continues to record the property at its original purchase price (less depreciation). While the FASB demands that they do this for external financial reporting, shouldn’t internal reports reflect the increase in wealth that you made possible? How will you make these arguments without clearly distinguishing between how events actually alter the wealth of your organization, and how that wealth is reported?

Conceptual and Operational Definitions

One reason we often forget to be accounting skeptics is that everyday language often hides the distinction between forms and shadows. Imagine that an accountant is asked:
Belief, Skepticism and Postmoderism in Accounting

“Is the cost of Nike’s brand-bolstering marketing campaign an asset or an expense?”

As posed, the question isn’t precise enough to be answered. Which of the following does the questioner actually mean?

“Is the cost of Nike’s brand-bolstering marketing campaign best viewed as an asset or an expense?”

“Is the cost of Nike’s brand-bolstering marketing campaign reported as an asset or an expense?”

To answer the first question, the accountant must refer to the conceptual definitions provided (courtesy of the FASB) in the essay “Elements of Wealth and Income”. This question is hardly trivial. On the one hand, “Just Do It!” is one of the most widely-known slogans in the world. Surely, the brand represents a source of “probable future economic benefits obtained or controlled [by Nike] as a result of past transactions or events.” On the other hand, surely some of the benefits arising from the marketing campaign are consumed by the end of the reporting period, so only a portion of its cost constitutes a future benefit at the moment the financial statements are compiled. How big is that portion? How would it be measured? If an event occurred that caused a change in that value (such as a high-profile scandal involving its key spokesman), how would the value of the asset be altered to reflect that change?

The second question is far easier to answer, because it is a question about Nike’s reporting choices, not the fundamental nature of the transaction. When Nike’s bookkeeper credited Accounts Payable to reflect the marketing spend, she debited either a balance sheet account or an expense account. If she debited a balance sheet account, Nike will report the campaign as an asset (and will probably amortize some of that amount each year). If she debited an expense account, Nike will report the campaign as an expense. The accountant is answering the second question by referring to a set of operational definitions, which classify elements of wealth and income by the operations undertaken by the bookkeeper (which accounts are debited and credited, and how they are reported).

For example, consider this pair of definitions for assets:

**Conceptual Definition:** Assets are probable future economic benefits obtained or controlled by a particular entity as a result of past transactions or events.

**Operational Definition:** Assets are items reported on the left side of the balance sheet.

Here is the analogous pair of definitions for expenses:

**Conceptual Definition:** Expenses are outflows or other using up of assets or incurrences of liabilities (or a combination of both) from delivering or producing goods, rendering services, or carrying out other activities that constitute the entity’s ongoing major or central operations.

**Operational Definition:** Expenses are items reported as components of Cost of Goods Sold or Selling, General, and Administrative Expenses on the income statement.
You won’t have much luck convincing people to adopt your preferred reporting practices unless you can distinguish clearly between the conceptual definitions that are guiding your recommendations and the operational definitions that describe the practices you are proposing. And you must help others make that distinction as well.
Part V: Allocating Overhead
The Mechanics of Overhead Application
The Mechanics of Overhead Application

Introduction

This essay walks through the mechanics and intuition involved in applying overhead to output in a normal absorption costing system, the most common type of costing system used in modern manufacturing and service firms.

The four key steps are as follows:

- Step 1: Calculating Predetermined Overhead Rates
- Step 2: Recording the Incurrence of Overhead
- Step 3: Applying Overhead to Cost Objects
- Step 4: Adjusting Cost of Goods Sold for Under- or Over-Applied Overhead

What Is Overhead?

Overhead costs are sacrifices of shared resources—resources consumed to achieve multiple goals, which therefore (in the terminology of cost accounting) must be allocated and applied to multiple cost objects, such as customer contracts, individual products, product lines, or departments.

Some of the most common shared resources are fixed costs. The fixed costs of owning a factory building and keeping it fit for use include rent, insurance, land taxes, lighting, climate control, water, and janitorial services. Every fixed cost is a shared resource, since (by the definition of “fixed cost”) the amount of the resources consumed does not change if one makes one more or one less unit of output. Variable costs can also be shared resources. For example, an auto-parts firm may use a great deal of energy to run its machines, and may use up many tools through wear and tear. These costs are variable because the more parts the firm creates, the more energy it uses and the more tools it wears out. However, they are shared resources because the firm probably has no direct way to identify which particular part was responsible for a particular component of the utility bills or tool replacements.

This essay focuses on recording one important class of overhead costs—those that are part of the cost of creating inventory. Recall from the essay “Elements of Wealth and Income” that the conceptual definition of inventory “includes assets that are held for sale to customers or for conversion into an asset that is held for sale to customers.” Costs that satisfy this definition are called inventorable costs or product costs and pass through Cost of Goods Sold when the inventory is sold to customers. Other overhead costs are simply the costs of maintaining ongoing operations; these we record in Selling, General, and Administrative Expenses, and we don’t address them in this essay.
Step 1: Calculating Predetermined Overhead Rates

The first step in accounting for (inventoriable) overhead is to calculate predetermined overhead rates. This requires the accountant to identify the set of shared resources to be included in the overhead pool, and to identify overhead drivers that can be measured separately for each cost object when it is produced.

Absorption costing systems include fixed costs, as well as variable costs, in the overhead pool. Common elements of an overhead pool are factory rent, janitorial services, building and machine maintenance, costs of information technology services, factory supervision, regulatory and licensing fees, engineering costs, taxes, and insurance. Keep in mind that the manufacturing overhead pool includes only costs associated with creating products or services to be delivered to customers. We place costs associated with sales, marketing, corporate activities, and general administration directly onto the income statement as expenses, in the Selling, General, and Administrative Expenses account.

Once an accountant has identified the resources to be included in the overhead pool, he or she must identify an overhead driver. An overhead driver must be easily measured for each cost object, and should have a reasonable causal relationship with overhead costs. For example, if electricity is the largest piece of the overhead pool, it makes sense to use machine hours as the overhead driver, since electricity usage probably increases with additional machine hours.

Some common overhead drivers are:

- Direct labor hours
- Direct labor dollars
- Direct material quantities (e.g., pounds, gallons, kilograms, liters)
- Direct material dollars
- Machine hours
- Billable hours
- Revenue dollars
- Customers served
- Units of production (e.g., basketballs produced)

Once you have identified the overhead driver, simply divide the projected amount of overhead incurred during the period (the overhead pool) by the total projected amount of the driver to be used during the period (the activity base).
Step 2: Recording the Incurrence of Overhead

Once the period has begun, the firm must record the incurrence of overhead for each pool. To do so, the firm credits the accounts that hold the shared resources being used. All of these accounts will reside on the balance sheet. If the firm is consuming resources it already possesses, the credits will be to asset accounts, such as Cash, Prepaid Expenses, Accumulated Depreciation, and Raw Materials Inventory. If the firm is borrowing to finance consumption, the credits will be to liability accounts, such as Accounts Payable or Wages Payable.

What account should be debited to balance the credits to resource accounts? Ideally, the firm would immediately debit the particular cost object that represents the purpose of incurring the cost. This isn’t possible for overhead, however, since (by definition) overhead represents the cost of resources shared by many cost objects. Instead, the firm debits a “Manufacturing Overhead Account” (MOH). For service firms, this might be called simply an “Overhead Account,” but to keep things simple while emphasizing the fact that we are addressing inventoriable costs, we will refer to MOH for any type of firm.

The MOH account is often called a suspense account, because it serves as a way station where costs sit while the firm waits in suspense to learn which cost objects they should be applied to. Note that it is an asset account (in the top left of the Queen’s Cross), even though no tangible item is associated with the dollars in the account.

Step 3: Applying Overhead to Cost Objects

As soon as the firm records the use of the overhead driver, it has enough information to transfer costs out of the Manufacturing Overhead Account into the Work-in-Process (WIP) Inventory account associated with the cost object responsible for the activity. The firm credits MOH and debits the appropriate WIP account for the same amount. This transfer is called cost assignment (also called cost application), and the account receiving the costs is the cost object.

The amount applied to a particular cost object is simply the product of (1) the predetermined overhead rate, and (2) the quantity of the overhead driver used by the cost object. Note that this calculation combines information from the firm’s budget with information from the firm’s records of actual activity. The budget provides the predetermined overhead rate (the projected size of the overhead pool divided by the projected activity base). The predetermined overhead rate must be based on projections because, as its name suggest, it must be determined before the period begins. Records of actual activity provide the quantity of driver use for the particular cost object to which overhead costs are being applied. This calculation of applied overhead is called normal costing because it is so commonly used.

The entry recording overhead application should be made as soon as you know the actual driver use for a cost object. Doing so allows the timely creation of a job-cost record, which summarizes all of the costs associated with a particular contract. Here is a sample job cost record from the solution to Exercise 3-29 of Ron Hilton’s Managerial Accounting textbook (9th edition, slightly modified). Note that it includes the cost of the resources directly consumed by the production of a lot of teddy bears (direct costs and direct...
labor), along with the overhead applied to the lot on the basis of direct labor hours. The record even includes supporting data: requisition numbers for raw materials used, and time card information for the direct labor hours. The job-cost record is the most important internal report of a job-order costing system, and its creation is the primary reason to allocate overhead from the MOH account to the WIP accounts associated with specific jobs.

<table>
<thead>
<tr>
<th>Job Number</th>
<th>Description</th>
<th>Date Started</th>
<th>Date Completed</th>
<th>Number of Units Completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>TB78</td>
<td>teddy bears</td>
<td>4/1</td>
<td>4/15</td>
<td>1,000</td>
</tr>
</tbody>
</table>

**Direct Material**

<table>
<thead>
<tr>
<th>Date</th>
<th>Requisition Number</th>
<th>Quantity</th>
<th>Unit Price</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/1</td>
<td>101</td>
<td>400</td>
<td>.80</td>
<td>$320</td>
</tr>
<tr>
<td>4/5</td>
<td>108</td>
<td>500</td>
<td>.30</td>
<td>150</td>
</tr>
</tbody>
</table>

**Direct Labor**

<table>
<thead>
<tr>
<th>Date</th>
<th>Time Card Number</th>
<th>Hours</th>
<th>Rate</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/1 – 4/8</td>
<td>BC22, BC23, BC24, CZ19</td>
<td>500</td>
<td>$12</td>
<td>$6,000</td>
</tr>
</tbody>
</table>

**Manufacturing Overhead**

<table>
<thead>
<tr>
<th>Date</th>
<th>Activity Base</th>
<th>Quantity</th>
<th>Application Rate</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/15</td>
<td>Direct-labor hours</td>
<td>500</td>
<td>$2</td>
<td>$1,000</td>
</tr>
</tbody>
</table>

**Cost Summary**

<table>
<thead>
<tr>
<th>Cost Item</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Direct Material</td>
<td>$ 470</td>
</tr>
<tr>
<td>Total Direct Labor</td>
<td>6,000</td>
</tr>
<tr>
<td>Total Manufacturing Overhead</td>
<td>1,000</td>
</tr>
<tr>
<td>Total Cost</td>
<td>$7,470</td>
</tr>
<tr>
<td>Unit Cost</td>
<td>$ 7.47</td>
</tr>
</tbody>
</table>

**Shipping Summary**

<table>
<thead>
<tr>
<th>Date</th>
<th>Units Shipped</th>
<th>Units Remaining In Inventory</th>
<th>Cost Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/30</td>
<td>700</td>
<td>300</td>
<td>$2,241*</td>
</tr>
</tbody>
</table>

Step 4: Adjusting Cost of Goods Sold for Under- or Over-Applied Overhead
The balance of the MOH account fluctuates throughout the period, and may have either a debit or credit balance, depending on the timing of resource consumption and driver use. These interim fluctuations are of little concern, because the MOH balance is not part of any important cost report.

The balance of the MOH account at the end of the period reflects errors in the initial projects used to calculate the predetermined overhead rate. If overhead actually incurred is less than overhead applied, the MOH account will end with a credit balance and the firm is said to have overapplied overhead. If overhead actually incurred is more than applied, the MOH account will end with a debit balance and the firm is said to have underapplied overhead.

An end-of-period balance in the MOH account would be misleading. Although MOH is a balance sheet account, an end-of-period balance does not reflect either an asset or a liability. If the MOH account has a debit balance, it does not reflect “probable future economic benefits” (the conceptual definition of an asset). It simply reflects that too little overhead was applied to WIP due to an estimation error. If the MOH account has a credit balance, it does not reflect “probable future sacrifices of economic benefits” (the conceptual definition of a liability). It simply reflects too much overhead applied to WIP due to an estimation error.

To avoid misleading readers of financial statements, firms compile their income statements after making an adjusting entry: they make whatever entry is necessary to reset the balance of MOH to zero, and balance that entry with a corresponding entry to Cost of Goods Sold. If overhead was underapplied, the firm needs to credit MOH (because it has a debit balance) and debit CGS. The resulting balance in CGS, called Adjusted Cost of Goods Sold, is higher than its balance before the adjustment, called Unadjusted Cost of Goods Sold, because the entry adjusts for the fact that not enough overhead was applied to inventory. If overhead was overapplied, the firm needs to debit MOH (because it has a credit balance) and credit CGS. In this case, adjusted Cost of Goods Sold is lower than Unadjusted Cost of Goods Sold because the entry adjusts for the fact that too much overhead was applied to inventory.

You might be wondering why the entire estimation error in the MOH account is used to adjust Cost of Goods Sold. After all, some of the over- or under-applied overhead is associated with cost objects still on the balance sheet (in Work-in-Process of Finished Goods Inventory). Firms rarely adjust inventory accounts for two reasons. The first is purely administrative. A firm may work on 10,000 different cost objects during a year, with 1,000 remaining in inventory at the time MOH is closed out. Even with computers, making all of those entries and verifying that they are correct would be no small task.

The second reason is that managers use job-cost records throughout the year as the basis for decisions on how to price jobs, and whether to take on new jobs. Adjusting WIP accounts would almost surely entail adjusting the job-cost report. That in turn would lead to memos such as the following:

“Virginia, I know you thought the cost of that job was $20,000 when you negotiated a price of $21,500 in February. Due to lower-than-normal volume in November, we have adjusted the cost of that job upward to $22,000. Be prepared to discuss your $500 loss at your next performance review.”
As you might guess, memos of this sort would not be part of good management practice. Instead, most firms simply correct for over- or under-applied overhead with a single adjusting entry to Cost of Goods Sold.
Introduction

A continuing theme in managerial reporting is that a reporting system should be matched to its organization’s needs. This essay describes the three primary goals of costing systems, and discusses how organizations choose costing systems that emphasize the goals most important to them. It begins with a definition of the rather slippery term, “cost.”

What is a cost?

A cost is a sacrifice, usually measured by the resources given up, to achieve a particular purpose.

- *Cost* is not a synonym for *expense*. An expense is a reduction in wealth incurred to support ongoing operations. When a power company buys coal, it has incurred a cost (sacrificing cash to get coal), but it has simply exchanged one asset for another and its wealth is unchanged. However, when it burns the coal, it incurs an expense because its wealth is reduced—and the cost of the coal consumed is the measure of that expense.

- *Cost* is also not a synonym for *expenditure*. An expenditure is a sacrifice of cash or a cash equivalent. Buying coal and burning coal both entail costs, but only buying it is an expenditure. Accountants often highlight this distinction by distinguishing resource spending (expenditures) from resource consumption (using up a resource that has already been purchased). Managers are often puzzled when they find that they have reduced their costs but their profits do not increase. This can happen if the managers have reduced their resource consumption, but not their resource spending. For example, they may use less machine time for each unit they make, but they may still spend exactly as much to lease their machines!

Accountants often describe costs in terms of the resource given up. If the resource is cash, the cost is called an out-of-pocket cost. The costs of sacrificing labor and materials to create inventory (unsurprisingly called “labor costs” and “materials costs”) are so essential to business operations that they are collectively called prime cost. Capacity costs arise when an entity uses up the time or space of a machine, building, or other large asset that can process or house only so much. For example, the cost of storing materials in only one-quarter of a warehouse sacrifices a limited resource (storage space).

Some costs consume resources that can’t be seen or touched. If the entity is sacrificing an opportunity to generate wealth, the cost is called an opportunity cost. If the entity is sacrificing competitive...
Different Costs for Different Purposes

In order to make informed decisions, businesses need to understand the costs associated with their operations. Opportunity costs arise when a firm has a resource that is limited but free, such as time; devoting time to one task sacrifices the opportunity to devote it to another. Proprietary costs arise when a firm must disclose information; the disclosure doesn't generate many out-of-pocket costs, but it may reveal trade secrets or, when reporting large profits, encourage competitors to enter a market.

Three Goals of Costing Systems

Costing systems seek to accomplish three goals:

- **Reporting Product and Period Costs.** Product costs are those incurred to provide the products and services sold to customers. Period costs are those incurred to maintain operations during the current period. Clear distinctions between these costs ensure that an organization does not show costs as assets (inventory) on their balance sheet when those costs provide no probable economic benefit after the period is over, which would inflate both assets and income.

- **Reporting Margins.** Organizations also need to know the cost associated with each individual source of revenue (such as each product or client). This information allows organizations to determine how much revenue they need each source to earn to make that particular activity profitable.

- **Reporting Efficiencies.** Finally, organizations must know whether they are providing products and services at the lowest possible cost. The more clearly the organization reports efficiencies and inefficiencies, the better they can identify opportunities to cut costs and increase profits.

No single accounting system can accomplish all three goals at once. This essay emphasizes the trade-offs between reporting margins and reporting efficiencies. The essay “The Challenge of Accounting for Fixed Costs” emphasizes the trade-off between distinguishing product and period costs and reporting margins.

Costing Needs of Job Shops, Process Shops, and Hybrid Shops

Businesses choose costing systems to report the most useful information for decision-making, given the nature of their production process and product market. A useful simplification is to distinguish between two extremes: job shops and process shops.

**The Job Shop**

A business that sells highly customized products or services is called a job shop, because each different contract to serve a customer is called a job. For example, contracts to repair a particular car, to remodel a particular kitchen, or to provide particular legal or consulting services to a client are all jobs, and these businesses would all be job shops.

Because job shops tailor their output to each customer, they typically enjoy enough product differentiation that they have the power to set their own prices. For example, competing proposals for a kitchen remodeling job probably differ in many ways. The greater degree of product differentiation allows job shops to charge high margins without losing too much of their customer base, because
customers can’t get an identical service from a competitor. If there were such a competitor, the remodeler would lose their price-setting power because they would be forced to match their competitor’s prices.

Job shops typically place high priority on reporting margins on a job-by-job basis (relative to reporting efficiencies), for several reasons. First, because they have the power to set prices, they can use reported margins to make better pricing decisions. Second, because every job is different, it can be difficult to compare how efficiently they completed a job, because they have no reliable standard for comparison. Third, even if they could complete a job slightly more efficiently, the variations from one job to another make it less likely that their improvement will produce further savings as they complete job after job.

To help them measure job-specific margins, job shops use job-order costing systems, which treat each job as the primary cost object to which costs are applied. Applying costs to jobs can be split into an easy part and a hard part.

- The easy part entails tracking the costs of resources used by one job and one job alone. For example, a window that is installed in my kitchen is installed in no other kitchen; therefore, the cost to the contractor of purchasing the window can be assigned directly to one job (my kitchen remodeling) and no other. The electricians wiring my kitchen are working on my kitchen for that time, and no other. Thus, their labor costs can be assigned directly to one job and no other. Such costs are called direct materials costs and direct labor costs, and are collectively referred to as direct or prime costs.

- The hard part entails applying the costs of shared resources that are used in many jobs. A contractor uses a truck and expensive tools to work on many different remodeling projects. How much of the cost of the truck should be applied to my kitchen remodeling job (as opposed to the many other jobs the contractor is working on)? What if one job requires extensive truck usage because it is far away and entails hauling a lot of material? What if the job uses expensive tools only a small amount of the time, relative to other jobs? Deciding how to allocate (divide up) such indirect costs or overhead is a major part of cost accountant’s job, and most cost accounting courses devote substantial time to exploring alternative methods of cost allocation.

The Process Shop

At the opposite end of the spectrum from the job shop is the process shop. While a prototypical job shop never does the same thing twice (every job is different), a prototypical process shop consists of different departments that do exactly the same thing over and over.

Because process shops engage in mass production, they tend to have very little differentiation from their competitors. In the extreme, a process shop is selling a commodity product, which by definition is a product with no product differentiation at all. In standard economic theory, commodity markets are perfectly competitive, which means that the producers are purely price-takers: they can sell as much as they can make at the market price, but they can’t sell anything for a penny more. (For example, farmers don’t get to set the prices of their produce; they simply take the going market rate.)
Different Costs for Different Purposes

Process shops typically place high priority on reporting efficiencies (relative to reporting margins), for several reasons. First, process shops have little power to set prices; even if they knew their margins, they couldn’t use that information very effectively. Second, the repetitive nature of process shops means they can easily assess the efficiency of their production process. Because a process shop does the same thing over and over, managers have a wealth of data to use in assessing how much a unit of output “should” cost, and to compare that to the actual cost they estimate. Moreover, saving a penny every time they engage in a process is very important for process shops, because they repeat that process so many times (and therefore save that penny so many times). Finally, process shops have such narrow margins that a small savings might mean the difference between profit and loss, and therefore between survival and bankruptcy.

To help them report cost efficiencies, process shops use process costing systems, often incorporated into standard costing systems. Process costing systems use the process (or more often, department) as the primary cost object to which costs are applied. Standard costing applies costs at what it should have cost to perform a process, rather than what it actually did cost. As with job shops, costing for process shops has an easy part and a hard part.

- The easy part is to calculate how much cost should be applied to each iteration of a process. Because the cost of doing the same thing over and over should be the same every time, the cost accountant can simply divide the total costs incurred in the department by the number of times it performed the activity. Consider a department that assembles 1,000 toys a month from purchased components, at a total cost of $20,000. If every toy is identical, the assembly cost can be estimated at $20,000/1,000 units = $20/unit assembled. There is no need to measure the time taken to assemble each unit, because there is no reason one assembly should cost more than another. There is also little reason to distinguish between direct labor and overhead in process shops—the two are simply summed up as a processing cost (or conversion cost) and allocated evenly among every unit.

- The hard part is to separate reporting of what a process should have cost from reporting of what it actually cost. Standard costing systems require firms to integrate their budgeting system with their cost accounting system, and to use a complex set of bookkeeping entries to report inefficiencies in a visible and understandable way. Because of their complexity, standard costing systems are discussed in later essays.

Mass Customization and Hybrid Shops

Many production processes can be characterized as mass customization businesses. As in job shops, output is customized to meet each customer’s needs. However, as in process shops, the production process is organized into separate departments such that each department does exactly the same thing over and over (mass production). For example, consider a producer of candied nuts. They have a Receiving department that sorts, cleans, and prepares purchased raw nuts, a Roasting department that roasts the nuts, and a Coating department that applies candy coatings.
While not every product passes through every department (some nuts are received and roasted but not coated, for example), within a department every nut is processed in exactly the same way. This departmental structure makes cost accounting nearly as simple as in a process shop. Within each department the nuts are so similar that they should all receive the same amount of cost. The cost of a particular job is simply the cost of its raw material, plus the processing costs of each department the job passes through. This form of costing is called operations costing.

Business that are large enough may find it worthwhile to reorganize their departmental structure to use operations costing—they simply create a “department” for each repetitive process. Such a reorganization makes sense when the costs of shared resources can be easily allocated to each of the newly formed departments, and when those departments do more or less the same thing repeatedly (as in a process shop).

**A Taxonomy of Costing Systems**

This essay introduces a handful of costing systems (job-order, process, and standard costing), but subsequent essays present many more, each designed to meet different organizational goals. The following page provides a brief reference guide to five key choices that organizations can make in designing their costing system. I present the reference guide here, in hopes that readers will find it easier to understand how the systems fit into the overall framework as they are introduced one by one.
## COSTING SYSTEM REFERENCE GUIDE

Every costing system reflects five distinct choices. Any combination of choices is possible, though some are more useful than others. (Few companies would choose to have a job-order standard costing system, for example.)

<table>
<thead>
<tr>
<th><strong>Primary cost object:</strong> The object to which costs are applied for the organization’s most important internal costing reports.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Job-order</strong> costing uses output objects, such as jobs, customers, and product lines, which are reported in job-cost records.</td>
</tr>
<tr>
<td><strong>Process costing</strong> uses process objects (such as departments, shifts, and assembly lines), whose costs are reported in departmental cost reports.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Activity Base:</strong> The denominator used to calculate predetermined overhead rates.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Projected production</strong> allows the firm to record all of its fixed costs in WIP and Unadjusted CGS; it is used in normal costing.</td>
</tr>
<tr>
<td><strong>Projected capacity</strong> helps the firm avoid a death spiral and keeps the firm from charging customers for the costs of excess capacity; it is used in capacity costing.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Pool Construction:</strong> How Overhead Pools Are Defined and Associated with Drivers.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simple costing</strong> uses only a single overhead pool with a single driver; it is used in businesses with little overhead or little variation across products.</td>
</tr>
<tr>
<td><strong>Two-stage costing</strong> uses multiple pools, but the driver for each one is a unit-level cost.</td>
</tr>
<tr>
<td><strong>Activity-based costing</strong> uses multiple pools, but has at least one driver above the unit level (e.g., at the batch or product level).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Product Cost Definition:</strong> Which costs are applied to inventory and recognized as expense in Cost of Goods Sold (treated as Product Costs), rather than being expensed immediately (treated as Period Costs).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Absorption costing</strong> systems treat both fixed and variable resources consumed in production as product costs.</td>
</tr>
<tr>
<td><strong>Variable costing</strong> systems treat only variable resources consumed in production as product costs.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Measurement of Inputs for Applied Costs:</strong> How costs applied to inventory are calculated.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Normal costing</strong> systems apply costs based on the inputs <em>actually</em> used in production. Direct costs are applied based on actual quantities of direct inputs used at the actual prices of those quantities. Indirect costs are applied based on the actual quantities of drivers used at the predetermined overhead rates.</td>
</tr>
<tr>
<td><strong>Standard costing</strong> systems apply costs based on the inputs that <em>should have been</em> used in production according to performance standards. Direct costs are applied based on standard quantities of direct inputs used at the standard prices of those quantities. Indirect costs are applied based on the standard quantities of drivers used at the (standard) predetermined overhead rates.</td>
</tr>
</tbody>
</table>
Improving Overhead Allocation Through Two-Stage Costing

Introduction

What distinguishes a good overhead allocation system from a poor one? Good systems report costs that more accurately reflect the shared resources consumed to produce each cost object (such as a product line, individual unit produced, or client contract). This ability to isolate margins allows an organization to sniff out the difference between profitable and unprofitable objects. It usually requires a system that divides overhead into multiple pools; the costs in each pool are allocated on the basis of a proxy measure that captures the true cause of overhead with as little measurement error as possible.

This essay demonstrates the structure and impact of such systems, after describing a single-driver system that serves as a basis for comparison.

Single-Driver Overhead Allocation

Consider a professional-services firm, such as a law, consulting, programming, or accounting firm. Such firms have substantial shared resources, including buildings, computers, secretarial staff, technical staff, human resources and training staff, and staff for regulatory compliance. How should these costs be allocated to individual client contracts? The key is to link the costs of shared resources that can’t be measured separately for each cost object to variables that can be.

The simplest method of overhead allocation is to link the dollar cost of the shared resource to a single measure. Assume that the firm expects building and occupancy costs for the firm to total $1 million over the year, and also expects to record 20,000 billable hours—hours worked by the professionals directly linked to a particular client contract (most professionals keep track of time spent working on each client contract, often down to five-minute increments).

The firm can then create a simple predetermined overhead rate by dividing the $1 million of projected overhead by the 20,000 projected billable hours (BH), resulting in an average rate of $50 of overhead per BH. The numerator (in this case $1 million) of the ratio is called the overhead pool, and the denominator (20,000 BH) is the activity base. The specific measure used to allocate overhead (in this case, billable hours) is called the overhead driver (although we often say “driver” for short).

The following diagram captures this very simple method of overhead application. The total overhead costs to be allocated are gathered in an overhead pool; these costs trickle down from the pool into each cost object at the predetermined overhead rate of $50 for every billable hour. Thus, a contract that
Improving Overhead Allocation Through Two-Stage Costing

requires 400 billable hours will be allocated $50 \times 400 \text{ BH} = 2,000$ of overhead cost. Note that this system is feasible only if the cost accounting system tracks the driver use for each individual cost object.

Simple Overhead Allocation

The Challenge and Importance of Capturing True Cost of Individual Cost Objects

The simple overhead allocation method described might be a reasonable means of reaching one of the goals of costing systems (as described in “Different Costs for Different Purposes”): it distinguishes clearly between product and period costs, as long as it includes the right costs in the overhead pool. The system provides almost no information about another goal, which is to report efficiencies, because it provides no baseline for comparing the cost of an object to what it should have cost. However, this essay focuses on the remaining goal: reporting the correct margin for each individual cost object.

Reporting accurate margins is an ambitious goal. One of the first lessons learned in financial accounting is that it is not too hard to measure the profitability of a firm over its entire life: just add up all of the distributions to shareholders (including upon dissolution), and subtract all of the contributions from investors. It is far more difficult to assess profitability over a given decade, much less a quarter or week. Similarly, it is far more difficult to assess the costs incurred by a department, much less a single product line, service, or job.

A useful starting point is to represent the causal model underlying an overhead allocation system, as shown in the Libby Boxes below. Note that driver use is simply a proxy measure for the activities that cause an object to consume shared resources. Similarly, the amount of overhead applied to a cost object is just a proxy measure for the true amount of shared resources it consumed. Both of these measures are probably quite noisy, so applied overhead is merely a shadow of the true cost, and a poor one at that.
Identifying a good proxy measure for the true cost of a driver helps managers determine the true profitability of specific contracts, but this is not the only benefit. It also conveys to managers the true costs of the resources they use. A predetermined overhead rate acts as a tax on the driver. For the professional-services firm we have been using as our example, choosing billable hours as an overhead driver makes each one $50 more expensive than it would be if we chose another overhead driver. Assuming that some manager in the firm is evaluated on the basis of cost control, overhead costs serve a decision-influencing function. People who want to emphasize the tax-like nature of an overhead rate may refer to it by the synonym burden rate. President Ronald Reagan famously said, “If you want to kill something, tax it!” Given that billable hours are the source of the firm’s revenue, it might not be the wisest thing to tax, because it will be borne either by customers (in the form of higher prices) or by account managers (reducing their reported margins, a primary basis for performance evaluation). It would be a particularly unwise thing to overtax. However, firms sometimes choose to have a driver bear more than its true burden to discourage managers from using it.

Hidden Differences and Hidden Drivers

An overhead allocation model with a single driver can suffer from (at least) two different problems. One is that the model assumes that every unit of the driver generates the same amount of overhead. However, this is probably far from true. For example, partners might be supported by five assistants, while associates are supported by only one. As a result, an hour billed by a senior partner should be allocated more overhead cost than an hour billed by an associate. Because the accounting system ignores these differences, it suffers from concealing a hidden difference—the same rate is applied to every unit of an existing driver, when the rates should be different.
Another possible flaw is that the model completely ignores other activities that drive overhead. Assume, for example, that some client contracts involve the use of extremely expensive computer hardware and software. In this case, overhead is driven not only by billable hours, but by hours of computer time. Because the accounting system ignores this other overhead driver, it suffers from concealing a hidden driver.

**Allocating Overhead with Multiple Overhead Pools**

Hidden differences and hidden drivers can be revealed by using a two-stage costing system that divides overhead into multiple pools, and then associates each pool with its own driver. Assume, for example, that the professional-services firm is most concerned about the hidden difference noted above—a partner’s billable hour generates more overhead than an associate’s billable hour. The accountant would work with management to identify how to split the $1 million of total overhead into two pools, by analyzing the costs of each professional’s support staff, use of office space, and so on. The calculations below assume that a total of $300,000 is assigned to the partners, and the remaining $700,000 is assigned to the associates. However, partners generate only 4,000 billable hours, while associates generate the remaining 16,000. The same calculation as before (dividing the dollars in each pool by its driver’s activity base) results in a much higher predetermined overhead rate for partners than for the associates: $75 of OH per partner hour, compared to $43.75 per associate hour. Note that the total activity base is unchanged, because we have simply divided the original activity base (20,000 billable hours) into two types (4,000 partner billable hours plus 16,000 associate billable hours).

Alternatively, assume that the firm wants to address the hidden driver problem—some clients demand expensive computer services that generate costs not incurred by other clients. In this case, the accountant would work with management to split the $1 million total overhead into two pools by analyzing which costs are associated with computer-intensive services and which with noncomputer...
services. The calculations below assume that $300,000 of overhead is directly associated with computer-intensive services (hardware, utilities, support, software purchases and licenses, etc.). The remaining $700,000 of overhead is associated with the billable hours from the professionals. Each of these pools is associated with its own driver. We must therefore create a new driver that was previously hidden; in this case, it is “technical staff hours,” because the technical staff does the computer-intensive work. Given 3,000 technical staff hours (TH), the overhead rate is $100/TH, while the overhead rate per billable hour drops to $35/BH. \textit{Note that the total activity base is now changed, because we have 3,000 technical staff hours in addition to the previously measured 20,000 billable hours.}

### Revealing a Hidden Driver

<table>
<thead>
<tr>
<th>OH Pool</th>
<th>Activity Base</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate OH Pool $1M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prof OH Pool $700K</td>
<td>$700K/20K BH = $35/BH</td>
<td>$35/BH</td>
</tr>
<tr>
<td>Tech OH Pool $300K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OH Pool/Activity Base $300K/3K TH = $100/TH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OH Applied to Cost Object = $35/BH x BH Used by Object + $100/TH x TH Used by Object</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Who Gains and Who Loses?

Changing overhead drivers invariably changes the reported performance of managers, who are typically responsible for the gross margins they can generate on the cost objects they produce and sell. How can we tell which cost objects will go up or down in total overhead cost? A useful starting point is the \textit{peanut-butter principle:}

\textbf{The Peanut-Butter Principle:} Overhead costs are like peanut butter. Changing overhead pools and drivers will spread it around differently across cost objects, but won’t change the total amount to be spread.

Imagine a firm in which every cost object uses exactly the same ratio of the various drivers. Since they are identical, they have to move in the same direction. They can’t all go up, and they can’t all go down, so following the peanut-butter principle, they must not change at all. This leads to the following \textit{Balancing Rule:}

\textbf{The Balancing Rule:} The overhead allocated to a cost object won’t change if the drivers used by the object are in exact proportion to the activity bases used to allocate those drivers.
Improving Overhead Allocation Through Two-Stage Costing

You can easily verify this under the three cost systems we have constructed for our professional-services firm. Overall, the firm has 4,000 partner hours and 16,000 associate hours, a ratio of 1:4. A contract with 100 partner hours and 400 associate hours (also a 1:4 ratio) would get $25,000 in overhead under the original system ($50/BH x 500 BH), and would also get $25,000 in overhead under the system that reveals the hidden difference in the cost of partner and associate hours ($75/BH x 100 BH + $43.75 AH x 400 AH). Similarly, the firm has 20,000 billable hours and 3,000 technical hours, a ratio of 20:3. A job with 200 billable hours and 30 technical hours (also a 20:3 ratio) would get $10,000 in overhead under the original system ($50/BH x 200 BH), and would also get $10,000 in overhead under the system that reveals the hidden difference in the cost of partner and associate hours ($35/BH x 500 BH + $100/ TH x 30 TH).

Jobs that deviate from the balancing point will increase or decrease in cost, depending on whether they use a disproportionately high or low amount of the driver that gets more expensive under the new system. In the original system, all billable hours were “taxed” the same $50. After revealing the hidden difference, partner hours were taxed $75 and associate hours were taxed only $43.75. As a result, contracts that rely heavily on partner hours (more than 1 partner hour for every 4 associate hours) will suffer from the tax increase on partner hours more than they benefit from the tax decrease on associate hours. Similarly, contracts that rely heavily on associate hours (more than 4 associate hours for every partner hour) will benefit from the tax decrease on associate hours more than they suffer from the tax increase on partner hours.

The analysis is similar when we change the accounting system to reveal a hidden driver. In the original system, the firm was taxing $50 for billable hours, but giving away technical staff time for free! After revealing the hidden driver, the firm was taxing each billable hour only $35, but adding a new tax of $30 for each technical staff hour. As a result, contracts that rely heavily on computer-intensive services (more than 3 technical hours for every billable hour) will suffer from the tax increase on those services more than they benefit from the tax decrease on billable hours. Similarly, contracts that include few computer-intensive services (less than 3 technical hours for every billable hour) will benefit from the tax decrease on billable hours more than they suffer from the tax increase on technical staff hours.

Unintended Cross Subsidies

Following the peanut-butter principle, a bad cost allocation system won’t make costs too high or too low across the board. Instead, some contracts, jobs, or products will be forced to pay too high a tax for the inputs they use a great deal of. In this way, hidden differences and hidden drivers force some cost objects to cross-subsidize others. Cross-subsidies can be a useful thing, if they are intended. A professional-services firm may choose to charge less for a partner’s time than it truly costs, because there is an advantage to having a client spend time with a high-profile partner. They may choose to charge less for computer-intensive services than they really cost, in order to build share in a growing market. However, bad cost allocation systems impose cross-subsidies that a firm does not intend, and often is simply not aware of. While intentional subsidies can be good strategy, costing based on ignorance has little to recommend it.
Bidding for Oil Rights

Hockey great Wayne Gretzky once remarked, “You miss 100 percent of the shots you don’t take.” Business often seems the same way. You make no profit from the jobs you don’t get, so you make all of your margins from the jobs you do get. Winning the bidding for a contract should be a cause for celebration, right?

Surprisingly, winning the bid often predictably harms profits. The hazards of winning an auction first came to light when oil exploration companies bid against one another to purchase drilling rights. Each company asked its best geologists to estimate the amount of oil in the plot up for auction, and each quite reasonably based its bid on their geologists’ estimate. Invariably, the company that won the auction ended up disappointed—the well did not produce as much as estimated, and the company had lost money by winning the auction. This result was termed the winners’ curse.

It didn’t take long for economists to explain this puzzling result. The economists argued that geologists were good enough at their jobs to be able to generate unbiased estimates of a field’s value. However, their estimates were still noisy. Sometimes they would overestimate or underestimate. As a result, the average estimate would be just about right, but the highest estimate was usually an overestimate (that’s why it was the highest). If every company bid based on their own geologists’ estimate, the highest bidder would bid too much. This effect is called adverse selection—the winners of the auction are selected for a reason that is adverse to their interest (in this case, they won because they overbid).

This analysis also suggested a simple solution: each bidder should enter a bid based, not on their geologists’ estimate alone, but also on the fact that they won the bidding, and therefore probably overestimated the value of the oil rights. Prospective buyers therefore need to shade their bids downward to secure a profit.
The Winners’ Curse from the Seller’s Perspective

Prices tended to be too high in the auction for oil rights because many firms were competing to buy from one seller. But businesses are often in the opposite situation, with many firms competing to sell to one buyer. Many construction firms compete to build a factory; many lawyers compete to argue a case; many software firms compete to develop a Web site. Each one must estimate the cost of satisfying the client. Who is likely to enter the lowest bid, and therefore win the contract? The one that underestimated its costs by the most! Like the oil-rights buyers, the “winning” seller often loses money on the job.

Common Values and Common Costs

The winner’s curse for sellers isn’t exactly the same as the winner’s curse for buyers. Bidding for oil fields is a special case, because the value of the oil is the same for everyone, and drilling costs are small enough and similar enough not to matter much. For a common value auction such as this, there is no reason to think the winners won because they had a competitive advantage.

The analogy to a common value auction for sellers would be a common cost auction—one in which no seller has a significant cost advantage over any other. In such a case, the sellers with the lowest estimate probably just underestimated their costs, and they would need to bid well above their estimate to turn a profit on the work. Like the oil-rights buyers, they would need to bid enough to earn their desired margin knowing not only their cost estimate, but the fact that as low bidder they probably underestimated their cost the most. They therefore need to shade their bids upward to ensure that the contract is profitable.

However, common cost auctions are rare. In an auction with many sellers, some usually have a competitive advantage over the others. They might have idle capacity they can use at little cost, or be particularly well suited to the exact type of work the buyer wants, or simply be more efficient at everything they do. This makes it even more difficult to determine what to bid.

Noise, Competition, Market Size, and Shading

How much should a seller shade a bid upward to account for the winners’ curse? The answer depends on three variables: how noisy the cost estimates are, how much competitive advantage (or disadvantage) the firm has, and how many other bidders are competing for the work.

- **More noise, more shading.** The key to the winner’s curse is that the firm can learn something about their own costs by looking at competitors’ bids. If the firm has perfect information about its costs, there is no need to shade the bid—nothing can be learned from the fact that the firm was the low bidder (and therefore the low cost estimator). The noisier the firm’s estimate of its own cost, the more the firm can learn from competitors’ bids.
- **More advantage, less shading.** A firm with a strong competitive advantage also has little to learn from the fact that they won the bidding. Even if their cost estimate is noisy, a low-cost
provider doesn’t learn much from the fact that they bid less than their competitors. However, if a high-cost provider places the low bid, watch out! They probably substantially underestimated their costs.

- **More sellers, more shading.** As the quote at the top of this essay indicates, exposure to the winner’s curse increases with the number of sellers. If a firm is competing against only one or two other suppliers, not much can be learned from knowing that one’s own bid is the lowest. But if there are 50 other sellers, it takes a pretty severe underestimate of costs to be the lowest bidder.

**Distinguishing Competitive Advantage from Bad Costing**

How does a firm determine whether it has a true competitive advantage or disadvantage, rather than having a bad costing system that simply under- or overestimates the costs of serving customers? This assessment is always tricky, but here are some helpful guidelines.

- **Reported costs are merely a shadow of true costs.** It is always worth reminding your colleagues that reported costs for individual cost objects, such as a contract proposal, are noisy proxy measures (Plato’s shadows) of true costs.
- **Reported costs are noisiest in firms with varied product lines.** Recall the peanut-butter principle, which compares overhead costs to peanut butter: different methods of allocating overhead can change where the costs are spread, but they don’t make total costs increase or decrease. A firm that makes only one type of product can’t easily over- or underestimate the product’s cost, because that would violate the peanut-butter principle. However, a firm with many highly varied product lines can spread too much cost to some and too little to others. The more varied the product lines, the more difficult it is to figure out where the costs should go.
- **Look for systematic variations in competitive advantage across product lines.** Many costing systems misestimate costs because they use too few pools or too few drivers. Adding pools and drivers can reveal hidden differences in the true cost of a single driver across different settings, can reveal hidden drivers that generate overhead, or both. However, you don’t need to change the costing system to determine whether these errors are creating the appearance of competitive advantage or disadvantage. Simply assess whether the firm’s competitive position seems to vary across jobs in a predictable way. For example, consider a professional-services firm that allocates all overhead on the basis of billable hours, but (because they use the same rate for all hours) hides the difference between high-overhead partner hours and low-overhead associate hours. The firm would probably find that the high-margin products all use a great many partner labor hours rather than associate labor hours. Such systematic variation suggests a costing problem, rather than a competitive problem.
- **Identify possible sources of competitive advantage and disadvantages and assess them directly.** Cost accounting systems aggregate highly detailed operational data into summary dollar estimates. If margins on a product line appear unusually high or low, hypothesize some reasons that your firm may actually have an advantage or disadvantage in serving customers efficiently. Are your production lines brand new or old and creaky? Are your workers particularly
well-trained and motivated, or disaffected and resistant to change? Do you have state-of-the-art reporting systems, or does management write all of their analyses on the backs of napkins? Talk directly to your operational staff to determine whether they believe operations are running smoothly.

• **Pay attention to how you are allocating the costs of unused capacity.** Finally, keep in mind that there is no single correct way to allocate the fixed costs of resources that are partly unused. Are you allocating those costs to products, rather than writing them off as costs of unused capacity? What might your competitors be doing? You can afford to take on work at prices below reported costs, if those reported costs include large allocations of fixed costs.
Part VI: Capacity and Surplus
Cost-Volume-Profit Analysis

A Simple Causal Model of Costs

Firms often call on accountants to predict the profitability of a new product or service based on very little information: selling prices, production costs, and the volume of business. Accountants approach this problem by constructing highly simplified causal models that link the volume of business to the cost incurred. The simplest causal model of costs, called the Cost-Volume-Profit model (CVP), is captured by this equation

\[ C = F + vQ \] (1)

where

- \( C \) = Cash expended for production (cost)
- \( F \) = Fixed costs that do not vary with the level of activity volume
- \( Q \) = The level of activity volume (usually the volume of a product or service provided)
- \( v \) = The variable cost for every unit of activity volume, also called the marginal cost

Like any model, the basic cost model dramatically simplifies the world to highlight certain distinctions and relationships. The most important of these are:

- All production results in the immediate payment of cash, according to the formula. The business never has to pay large amounts up front to buy an asset that will last for many years.
- Fixed costs (\( F \)) are exactly the same regardless of the quantity of production. In reality, few costs behave this way, but it can be a good approximation in a limited range of production, called the relevant range.
- Variable costs increase exactly in direct proportion (\( v \)) to the quantity of production. In reality few costs behave this way, but it can be a good approximation in the relevant range.
- The quantity of production can be measured with a single variable (\( Q \)). This is rarely true, but even if we make multiple products we can use this model by assuming a constant product mix. \( Q \) then reflects a weighted average of the variable costs of each of the products we make.
A calculation of profit requires a similarly simplified formula for revenue,

\[ R = pQ \] (2)

where

- \( R \) = Total revenue received in cash at the moment of sale
- \( p \) = Cash received per unit sold

Assuming that the quantity produced is instantly sold, so that \( Q \) is the same in both equations (1) and (2), the equations can be combined to yield

\[ \pi = R - C = pQ - F - vQ = (p-v)Q - F \] (3)

where \( \pi \) (pronounced “pi”) denotes the firm’s profit.

Breakeven Analysis

Though very simple, equation (3) is extremely useful for analyzing relationships among cost, volume, and profit. Most marketing students learn one of these applications: breakeven analysis. This application assumes values of \( F, p, \) and \( v \) and calculates the level of \( Q \) (called \( Q_{\text{BE}} \)) required for profit to be 0. Solving the equation \((p - v)Q_{\text{BE}} - F = 0\) with a little 8th-grade algebra yields the equation

\[ Q_{\text{BE}} = F/(p - v). \] (4)

The quantity \((p - v)\) is referred to as the contribution margin (CM), because each unit produced and sold contributes \( p - v \) toward paying off fixed costs.

Cost-Volume-Profit Analysis

Breakeven analysis may be a staple of introductory marketing, but accountants are called on to apply the CVP model in many other ways. The simplest variations entail solving for a variable other than \( Q \) to hit a target level of profit. For example, a firm may have a good sense of the quantity they can produce and sell, and want to know the breakeven level of sales price, variable cost, or contribution margin. Recall from 8th-grade algebra that we can only solve for one unknown variable when we have only one equation. The accountant must therefore make assumptions about the other variables to get a unique answer (for example, assuming values of \( p, F, \) and \( Q \) to solve for the breakeven \( v \)—the variable cost per unit that will allow the firm to break even given the sales price, fixed costs, and volume).

A more challenging application of CVP analysis helps firms evaluate decisions that alter the mix of fixed and variable costs. Firms can often trade one type of cost for the other; the mix of fixed and variable costs is called the cost structure. Investments in automation increase fixed costs but reduce variable costs. Outsourcing of production reduces fixed costs but increases variable costs. The wisdom of such investment or disinvestment depends on comparing two versions of equation (1). Assume that the firm currently faces fixed cost \( F \) and variable cost \( v \), and could invest \( \Delta F \) to get a reduction in variable costs of
\( \Delta v \). (The symbol \( \Delta \) is pronounced “delta” and is commonly used in mathematics to refer to a change.)

The two cost functions are now:

\[
\text{No change:} \quad TC = F + vQ \quad (1)
\]

\[
\text{Change:} \quad TC_\Delta = F + \Delta F + (v - \Delta v)Q \quad (1a)
\]

Taking the difference between the two possibilities, we have

\[
TC_\Delta - TC = \Delta F - \Delta vQ. \quad (1\Delta)
\]

Intuitively, a firm should make the investment if the increase in fixed costs is small relative to the total savings from reducing variable cost per unit. That total will be very large if production volume is very high. Thus, automation is a good strategy for high-volume businesses. For the same reason, outsourcing (which reduces fixed costs but increases variable costs) tends to be a good strategy for low-volume businesses.

If this analysis seems easy, it is because we assumed that we could identify the two changes in cost (\( \Delta F \) and \( \Delta V \)). This is, of course, extremely difficult in real-world settings. How much will fixed costs actually increase if we invest in automation? What fixed costs can we avoid by outsourcing? How much will our variable costs actually change in either scenario? To address questions such as these, it is essential to identify the incremental costs or relevant costs (synonyms that refer to costs that differ across two scenarios).

Even in the classroom, many students find this analysis difficult because they fail to make the right comparison! Students familiar with breakeven analysis from introductory marketing often want to plug equation 1\( \Delta \) into equation (3) and solve to get the breakeven quantity for equation (4). However, this simply answers the question, “Will I break even with my new cost structure?” In fact, the question demanding an answer is, “Will I make more profit with my new cost structure than with my old cost structure?” Since the revenue equation is unchanged, we need only ask which structure gives us lower costs.

**Operational Leverage**

Changing a firm’s cost structure affects not just the level of profits, but the sensitivity of profit to changes in volume. This sensitivity is called operational leverage (OL). A simple measure of operational leverage is the ratio of fixed costs to total costs:

\[
OL = \frac{F}{F + vQ} \quad (5)
\]

A firm with high fixed costs and relatively low variable costs (perhaps because they are heavily automated) has high operational leverage, and their profit will be very sensitive to volume. When volume is low, they must still pay their high fixed costs. When volume is high, their fixed costs don’t
increase, and their low variable costs allow them to retain a very high contribution margin from every extra unit they sell.

The operational leverage factor (OLF) is a ratio that captures this sensitivity even more directly:

\[
OLF = \frac{\text{Total CM}}{\text{Profit}}
\]  

OLF is a useful ratio because the OLF times the percentage change in volume gives the percentage change in profit. For example, assume that CM = 1, F = 100, and Q = 300. Remember that CM = (p – v), so, CM = 300, and \( \pi = \text{CM} \times Q - 100 = 200 \). The results in OLF = 300/200 = 1.5. Now assume that Q increases by 10%, from 300 to 330. Since fixed costs don’t change, the increase in profit rises from 200 to 230, a 15% increase (or 1.5 x 10%). One shortcoming of OLF is that it isn’t nearly as stable as OL—a change in volume will cause dramatic changes in profit, and therefore dramatic changes in OLF, especially if profit is close to 0. And if profit is negative, the OLF isn’t a very meaningful number at all. Thus, most analyses of cost structures look at both OLF and OL.
Introduction

The Cost-Volume-Profit model provides an extremely simplistic description of how costs behave: some are perfectly fixed and others are perfectly variable. Despite this simplicity, the presence of fixed costs poses a serious challenge to cost accountants. The heart of the problem is that fixed costs make it difficult (if not impossible) to create a single system that accomplishes both of the first two goals of cost accounting laid out in the essay “Different Costs for Different Purposes”:

- Distinguishing Product and Period Costs
- Reporting Margins

To meet the first goal, accounting systems include fixed costs in overhead pools that are applied to cost objects through one or more overhead drivers; those fixed costs are held inventory the balance sheet until the objects are sold, because the resources consumed are not lost, but simply transformed into another asset (inventory). Such systems are called full costing systems (because they include all overhead costs, fixed as well as variable) or absorption costing systems (because cost objects “absorb” the fixed costs). Unfortunately, absorption costing systems overstate the marginal costs of production, because they make it seem as if more activity generates more fixed costs; thus, they fail to meet the second goal of accurately capturing the margin associated with each individual product or service sold. The alternative is to allocate only variable costs to each cost object. Variable costing systems can meet the second goal by accurately capturing marginal costs of production, but they fail to meet the first goal because they understate aggregate costs. As promised in the first pages of these essays, no system is perfect!

This essay explains in more detail why absorption costing poorly reflects the true costs of individual cost objects, and then discusses the types of decision errors this failure can induce. It then explains the variable-costing alternative, and the shortcomings of that alternative. Finally, the essay discusses the benefits and challenges of using both types of systems in a single firm.

Why Absorption Costing Poorly Reflects True Costs of Individual Cost Objects

Absorption costing does a fine job of capturing all product costs, but it does a poor job of reflecting the true cost of any individual cost object. The reason is that absorption costing reports fixed overhead as if it were variable. Remember that the basic mechanic of absorption costing is to calculate a predetermined overhead rate by dividing all the overhead in an overhead pool (fixed and variable) by the projected activity base of some driver. Every additional use of the driver takes a little of the fixed
overhead out of the overhead account and adds it to inventory. So the perceived incremental increase in reported cost for engaging in more activity includes both the variable cost and a portion of the fixed costs.

**Costing Errors Lead to Decision Errors**

Treating fixed costs like variable costs can lead managers to three related decision errors: mistakenly declining business opportunities that have a positive contribution margin, suppressing use of the driver used to allocate overhead, and producing units they cannot sell in the current period.

1. **Declining business opportunities with a positive contribution margin**

Consider a decision to take one more job for a price of \( p \). This is a decision to work on the margin, because, like any marginal analysis (or incremental analysis), the firm is making only a very small change in activity. Traditional CVP analysis focuses appropriately on the contribution margin (price minus variable cost). If the contribution margin is positive, taking the additional job increases revenue more than it increases cost. However, an absorption costing system includes fixed as well as variable costs in the reported cost of the product. What are managers likely to do when the contribution margin for a new job is positive (so taking it will increase profit), but the reported gross margin for the job is negative (because reported cost includes the job’s allocation of fixed costs)?

The decision-facilitating function of reported costs will suggest to the manager that taking the job would be unwise. Worse yet, the decision-influencing function of reported costs is likely to encourage the manager to decline the job. After all, many managers are evaluated and compensated on the basis of the reported margins for individual jobs, product lines, or other cost objects. But the overapplication of fixed overhead generated by the job (because the fixed costs are applied but not incurred) doesn’t show up in the reported margin for that specific job—it is included in the aggregate adjustment to Cost of Goods Sold when the MOH account is zeroed out at the end of the year.

A number of good strategic reasons exist to decline work that has a positive contribution margin but fails to cover its share of fixed costs. For example:

- If every job is priced at that level, the firm cannot make a profit overall.
- The firm may not want to make a pricing exception for one customer, lest other customers demand similar treatment.
- The firm may want to maintain high prices to maintain a high reputation for quality.

However, declining a job because a manager thinks it will reduce income, when it will actually increase income, is nothing more than a mistake, and one often caused directly by absorption costing systems.

2. **Underusing the driver used to allocate overhead**

Absorption costing systems place the burden of fixed costs on specific overhead drivers. In the simple case of a single driver, that driver is effectively taxed not only for the variable costs truly incurred if the firm uses more of the driver, but also for fixed costs incurred whether the driver is used or not. If the
fixed costs reflect an asset with unlimited capacity (such as a liquor license), depressing demand for the overhead driver is clearly unhelpful. Assume that a casino allocates part of the cost of a liquor license to every drink they give away for free to frequent (and lucrative) card players. The absorption costing system would tell managers that every free drink costs more than it actually does, possibly leading managers to be overly stingy.

The analysis is slightly more complex for fixed costs that provide limited capacity (so-called capacity costs). Assume, for example, that a firm allocates both fixed and variable overhead on the basis of machine hours. Including fixed costs in the overhead pool will reduce the consumption of a limited resource (machine time). However, reduced resource consumption does not increase profit until it translates into reduced resource spending. If the firm can reduce machine time enough to cut the number of machines it owns or leases, profit increases. But if they simply increase the idle capacity of their machines—the amount of time they sit unused—reducing resource consumption has no effect on the bottom line.

3. **Producing units they cannot sell in the current period**

The final dysfunctional effect of absorption costing is that it may influence managers to produce units that they cannot sell in the current period. The effect is indirect, because it operates through the adjustment to Cost of Goods Sold to close out over- or underapplied overhead. Before diving into the flow of costs in an absorption costing system, it is useful to highlight the basic intuition:

A dollar of cost can show up either on the balance sheet (as inventory) or the income statement (as Cost of Goods Sold). It can’t be in both places at once. Producing goods that aren’t sold puts more fixed cost on the balance sheet, and therefore puts less fixed cost on the income statement, and makes income look higher.

Note that this effect reverses as soon as the inventory is sold, and the fixed costs sitting in inventory are transferred to Cost of Goods Sold. Nevertheless, producing goods you can’t sell increases income in the short term, so absorption costing systems encourage managers to run their operations full-tilt at the end of the year to build up inventory (a decision-influencing effect). This behavior is exactly the opposite of what experts in operations management and finance recommend, because inventory costs money to track and store, and it ties up vital working capital. The problem is sufficiently important that Eli Goldratt devotes much of his book, *The Goal* (used in many operations management courses), to criticizing cost accountants for encouraging managers to use capacity to its fullest extent, regardless of the benefits.

**Variable Costing: An Imperfect Alternative**

How could accountants be so misguided as to use absorption costing, despite all of its problems? The answer, as in so many cases of apparently suboptimal behavior, is that the alternative has problems of its own.
The Challenge of Accounting for Fixed Costs

The most common alternative to absorption costing is variable costing (sometimes called direct costing). Variable costing differs from absorption costing in one very simple way: it treats fixed costs not as product costs, but as period costs—much as accounting systems treat Selling, General, and Administrative expenses. Under variable costing, fixed costs are not transferred into manufacturing overhead, do not flow into FG, and do not appear in CGS when items are sold. Instead, fixed costs are immediately debited to an expense account, Fixed Manufacturing Overhead, on the income statement.

The key benefit of this approach is that variable costing treats fixed costs as fixed. Instead of emphasizing the distinction between product costs (those that pass through inventory and are expensed only when items are sold) and period costs (costs incurred simply to permit the firm to administer its operation and sell its goods), a variable costing income statement emphasizes the distinction between variable and fixed costs. Variable costs, whether they are product or period costs, are subtracted from revenue to report contribution margin. Fixed costs, whether they are product or period costs, are then subtracted from contribution margin to calculate net income.

### Absorption Costing Income Statement

<table>
<thead>
<tr>
<th>Item</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue</td>
<td>XXXXX</td>
</tr>
<tr>
<td>Unadjusted Cost of Goods Sold (Fixed &amp; Variable)</td>
<td>&lt;XXXX&gt;</td>
</tr>
<tr>
<td>Adjustments to Cost of Goods Sold</td>
<td></td>
</tr>
<tr>
<td>Incurred Overhead</td>
<td>Xxx</td>
</tr>
<tr>
<td>Applied Overhead</td>
<td>Xxx</td>
</tr>
<tr>
<td>Adjusted Cost of Goods Sold</td>
<td>&lt;XXXX&gt;</td>
</tr>
<tr>
<td>Gross Margin</td>
<td>XXXXX</td>
</tr>
<tr>
<td>Selling, General, &amp; Admin Expenses (Fixed &amp; Variable)</td>
<td>&lt;XXXX&gt;</td>
</tr>
<tr>
<td>Net Income</td>
<td>XXXXX</td>
</tr>
</tbody>
</table>

### Variable Costing Income Statement

<table>
<thead>
<tr>
<th>Item</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue</td>
<td>XXXXX</td>
</tr>
<tr>
<td>Variable Costs of Production and SG&amp;A</td>
<td>&lt;XXXX&gt;</td>
</tr>
<tr>
<td>Adjustments to Variable Production Costs</td>
<td></td>
</tr>
<tr>
<td>Incurred Variable Overhead</td>
<td>Xxx</td>
</tr>
<tr>
<td>Applied Variable Overhead</td>
<td>Xxx</td>
</tr>
<tr>
<td>Adjusted Variable Expenses</td>
<td>&lt;XXXX&gt;</td>
</tr>
<tr>
<td>Contribution Margin</td>
<td>XXXXX</td>
</tr>
<tr>
<td>Fixed Costs of Production and SG&amp;A</td>
<td>&lt;XXXX&gt;</td>
</tr>
<tr>
<td>Net Income</td>
<td>XXXXX</td>
</tr>
</tbody>
</table>

The variable costing income statement illustrates how well variable costing addresses the three decision errors likely to arise under absorption costing:
The Challenge of Accounting for Fixed Costs

- Firms are unlikely to mistakenly decline business opportunities that have a positive contribution margin, because the contribution margin is clearly reported.
- Firms are unlikely to underuse the driver used to allocate overhead, because that driver is not burdened with the costs of fixed overhead.
- Firms are unlikely to produce items they can’t sell, because doing so has no income-increasing effect.

Shortcomings of Variable Costing

So why doesn’t every firm use variable costing instead of absorption costing? Because variable costing fails to meet the first goal of cost accounting: to distinguish appropriately between product and period costs! More generally, while variable costing does a better job supporting decisions at the margin, it does a worse job supporting decisions in the aggregate. The three types of mistakes managers might make with a variable costing system are essentially the flip side of those in an absorption costing system:

1. Because variable costing systems misrepresent fixed costs of production as period costs, they overstate overall margins, and may lead managers to be satisfied with prices that are too low to cover all of their fixed costs (because the margins seem so high on each individual item).
2. Misreporting fixed production costs as period costs also misstates income, because it understates the value of inventory on the balance sheet, and understates costs “incurred because the entity has parted with goods or services conveyed to customers as part of ongoing operations” (the definition of Cost of Goods Sold). These misstatements are sufficiently misleading that variable costing is prohibited for both external reporting (GAAP) and tax reporting in most jurisdictions.
3. Variable costing may lead firms to overuse of fixed resources, because managers are not charged for consuming them. This makes sense for marginal decisions, but the aggregate result can be quite damaging.

Maintaining Multiple Systems

One solution to the problem of fixed costs is to maintain multiple systems. Such an approach presents both administrative and behavioral challenges. From an administrative perspective, a dual system is complex and costly. The system must be designed to track information at a more refined level. To generate variable costing statements, the system must split both MOH and SG&A into fixed and variable components. To generate absorption costing statements, the system must generate predetermined fixed overhead rates and allocate fixed overhead to products. The system must also be designed to provide the appropriate statement to the appropriate decision-makers at the appropriate time.

From a behavioral perspective, managers must be trained to understand both types of statements, and why they are being provided with one rather than another. Even if accountants can survive this challenge, they face a greater one: most managers view financial statements as truth itself, not as shadows of truth. Providing multiple statements can lead managers to discount the validity of any accounting report, thinking that statements are nothing but a reflection of “those creative accountants.”
The Death Spiral and Capacity Costing

Introduction

The Cost-Volume-Profit model (\(C = F + vQ\)) assumes that the incremental cost of producing one more unit is the variable cost, \(v\). But absorption costing systems report that the incremental cost of producing one more unit is \(v + \text{unit's share of fixed overhead}\). Such systems allocate the fixed overhead \(F\) across all of the units the firm expects to produce, so that the reported incremental cost of the product is \(v + F/Q^*\), where \(Q^*\) is the expected level of activity, also called the denominator volume. The lower the activity base, the higher the portion of fixed costs absorbed by each unit of activity, and the higher the reported per-unit cost. The dependence of reported cost on the activity base can easily lead to decision errors, the most important of which is the Death Spiral.

The Death Spiral

Consider a firm with many product lines that all share the same set of resources in the same per-unit quantities, but that bring in different revenue per unit. What should the firm do if the least-profitable product line has a negative gross margin (reported costs are greater than revenue)?

Let’s call the negative-margin product line the FAILwidget. If the FAILwidget doesn’t play some strategic role for the firm, such as being a loss leader, managers will be very tempted to eliminate it. Readers comfortable with the CVP model will see quickly that this may not be a wise move. Assuming revenue of \(p\), the FAILwidget brings in a contribution margin of \(p - v\) for each unit produced and sold. It is entirely possible that the contribution margin is positive, but not large enough to offset the allocated fixed cost \(F/Q^*\). However, eliminating the FAILwidget won’t actually reduce those costs, because fixed costs reflect shared resources that are still required by other product lines. Instead, the firm’s profit will drop by the amount of the lost contribution margin.

If the firm never alters its activity base, eliminating the FAILwidget will change revenue, Unadjusted Cost of Goods Sold, and the adjustment to Cost of Goods Sold. Revenue will fall by \(pQ_{FW}\) and unadjusted Cost of Goods Sold will fall by \(Q_{FW}(v + F/Q^*)\), where \(Q_{FW}\) is the quantity of the overhead driver no longer used because the FAILwidget product line has been eliminated. However, because fixed costs, \(F\), have not changed, the adjustment to Cost of Goods Sold must rise by \(Q_{FW}F/Q^*\) to make up for the fixed costs incurred but no longer applied.

However, eventually the firm is likely to adjust its activity base to reflect the pared-down product line. At this point, the effect of eliminating the FAILwidget is more sinister. Instead of reporting the unapplied fixed costs as an adjustment to Cost of Goods Sold, the firm recalculates its overhead rate using a
smaller activity base, applying more fixed costs to each unit. The increased fixed cost allocation reduces the gross margin of the remaining products, and the least profitable of those—call it “NEXTwidget”—is at risk of elimination. Eliminating NEXTwidget further increases the costs of the firm’s remaining product lines. This iterative process is called the Death Spiral, because each elimination further increases the reported costs of the remaining product lines, until the firm’s total contribution margin fails to cover fixed costs, and the firm can no longer operate profitably. The Death Spiral is often accelerated by the link among reported costs, pricing, and demand. Firms are likely to raise prices when their reported costs increase, causing a further decline in demand, and thus a further reduction in the firm’s activity base.

Avoiding the Death Spiral

To see how firms can avoid (or at least mitigate) a Death Spiral, it is useful to identify four contributing factors in a typical sequence. First, customers don’t provide the firm with sufficient revenue to allow its contribution margins to offset its fixed costs and generate the firm’s target profit margin. Slack demand leads the firm to eliminate its worst-performing product lines, and subsequently lower its activity base to reflect lower productivity. The firm then increases prices to achieve its target profit margin on products with higher reported costs. The higher prices suppress demand for the remaining products, and the cycle begins again. In theory, the Death Spiral could occur without price increases and the resulting decline in demand. However, these effects are common, and exacerbate the spiral.
What should the firm do differently? Here are three recommendations:

1. **Don’t eliminate product lines just because reported margins are too low.**

   Reported costs in absorption costing systems estimate the total resource costs dedicated to production. It’s easy to be misled into thinking those costs will disappear if the firm eliminates the product line. But remember that reported costs measure the costs of resource *consumption*, not the costs of resource *spending*. Eliminating a product line may reduce demands for machine time, factory space, and labor, but if the firm doesn’t actually reduce their spending on machines, facilities, and labor, total spending—and therefore total costs—won’t decline. If the firm never eliminates product lines with positive contribution margins, it will never initiate a Death Spiral.

   Product-line decisions require a more detailed analysis that distinguishes between *avoidable fixed costs* and *unavoidable fixed costs*. Avoidable costs are those that will be eliminated along with the product line. Perhaps the firm can sell some machinery, pare down support services such as information technology, reduce the number of indirect laborers, and move to a smaller facility. Unavoidable costs are those that will remain—and be borne by remaining product lines—even if you eliminate the product line. Note that neither variable nor absorption costing systems distinguish between avoidable and unavoidable costs, so product-line decisions require a special analysis based on the specific decision being considered.

2. **Be careful when linking prices to reported costs.**

   Marketers and economists often downplay the role of reported costs in setting prices, arguing that prices should reflect customers’ willingness to pay. *Cost-based pricing* is popular because firms rarely have good information about customers’ willingness to pay. They therefore use reported costs—usually from an absorption costing system—as a proxy measure for market demand. Assuming there is an active market for a product, total reported cost (plus some target margin) is often a good estimate of the customers’ willingness to pay—otherwise the market wouldn’t be active!

   However, a major change in reported costs due to a lower activity base isn’t a reflection of customers’ demands or competitors’ costs. Firms should be wary of basing prices on fixed costs that simply reflect the cost of unused capacity, rather than the cost of serving customers.

3. **Don’t charge customers for capacity they don’t demand.**

   Costing systems drive the Death Spiral by allocating all fixed costs to the products and services sold to customers. This forces product managers to suffer lower margins or raise prices. One way to avoid Death Spirals is to use an alternative cost accounting system, called *capacity costing*, which avoids charging customers for unused capacity by excluding it from product costs in the first place.

**Introduction to Capacity Costing**
The Death Spiral and Capacity Costing

Until now, all of the costing systems we have seen use projected production as their activity base (the denominator used to calculate predetermined overhead rates). Eliminating a product line in such systems increases the costs of the remaining products, because firms reduce their activity base to reflect decreased production—dividing fixed overhead by a smaller activity base increases the amount allocated to each unit produced. This practice satisfies the accountant’s goal of accurately distinguishing product and period costs, because if production is close to projections there is little over- or underapplied overhead. However, it misrepresents the margins earned by each unit produced, by effectively making customers pay the cost of units that are not produced due to lack of demand.

If some fixed costs of production are not applied to output, fixed overhead will be underapplied. Capacity costing systems report underapplied overhead due to low production volume as a separate adjustment to Cost of Goods Sold, often called Cost of Unused Capacity, as in this report:

### Capacity Costing Income Statement

<table>
<thead>
<tr>
<th>Description</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue</td>
<td>XXXXX</td>
</tr>
<tr>
<td>Unadjusted Cost of Goods Sold</td>
<td>&lt;XXXX&gt;</td>
</tr>
<tr>
<td>Adjustments to Cost of Goods Sold</td>
<td></td>
</tr>
<tr>
<td>Costs of Unused Capacity</td>
<td>xxx</td>
</tr>
<tr>
<td>Other Under- or Overapplied OH</td>
<td>xxx</td>
</tr>
<tr>
<td>Adjusted Cost of Goods Sold</td>
<td>&lt;XXXX&gt;</td>
</tr>
<tr>
<td>Gross Margin</td>
<td>XXXXX</td>
</tr>
<tr>
<td>Selling, General, &amp; Administrative Expenses</td>
<td>&lt;XXXX&gt;</td>
</tr>
<tr>
<td>Net Income</td>
<td>XXXXX</td>
</tr>
</tbody>
</table>

Capacity costing avoids punishing remaining customers for the costs that would have been applied to lost customers, and draws attention to the firm’s success in limiting resource spending. If the firm reduces resource consumption without reducing resource spending, the difference appears as the Cost of Unused Capacity right as its own expense line—a powerful attention-directing and decision-influencing device!

**The Two Commandments of Capacity Costing**

Capacity costing forces firms to decide how much of the costs of capacity they should allocate to the costs of serving customers, and how much to the managers deemed responsible for unused capacity. Successful implementations of capacity costing follow two basic commandments:

1. Thou shalt charge the costs of unused capacity to those who demand it.
2. Thou shalt not charge the costs of excess capacity to customers.

The first commandment forces firms to think carefully about the different components of unused capacity (also called idle capacity), and to assign the costs to the people who demand that the firm
maintain capacity they aren’t using. Note that “those who demand capacity” could refer either to customers or to managers inside the firm. Here are some common forms of idle capacity, and how they should be charged according to Commandment 1:

- **Surge Capacity.** Surge capacity is the component of capacity left idle because of variations in customer demand. Ice cream factories have idle capacity in the winter because they need it to satisfy high demand in the summer. This capacity is idle because customers demand ice cream when they want it, and following commandment 1, its costs should be included in product costs.

- **Buffer Capacity.** Buffer capacity is the component of capacity left idle because of variations in output. Complex production processes with variable yields maintain idle capacity so that they can still meet customer demand on days when yields are low. This capacity is idle because customers demand on-time delivery even though the production process is variable, and following commandment 1, its costs should be included in product costs.

- **Strategic Capacity.** Strategic capacity is idle because some manager has made a strategic decision to invest in capacity that cannot be put to immediate use. A firm may invest in a 3-D printer that it uses only occasionally, to position itself for rapid growth should demand for 3-D prototyping take off. The manager who made the strategic decision demanded this excess capacity, and it should be charged to him or her directly, rather than included as a product cost.

- **Contractual Capacity.** Contractual capacity is idle because the firm has agreed to maintain idle capacity that it can quickly put to use if a customer wants it. For example, a car manufacturer may sign a contract with a government to maintain a larger plant than needed to serve current customers, in case the government demands rapid production during wartime. The costs of such capacity should be allocated to the specific party to the contract, not to other customers.

Even after accounting for all of these causes for idle capacity, the firm might have additional idle capacity because customers are simply not demanding as much activity as the firm can provide. Idle capacity due to a lack of demand (and usually due to a decline in demand) is called “excess capacity.” Commandment 2 emphasizes that excess capacity should not be charged to customers, because raising prices in the face of low demand is a sure path to even lower demand in the future, accelerating the Death Spiral:

- **Excess Capacity.** Excess capacity is the capacity that is idle because no one is demanding it. Commandment 2 emphasizes that customers should not be charged for the capacity no one wants to use. This will simply suppress demand further, driving the firm directly into a Death Spiral.

**Adjusting the Activity Base to Follow the Commandments of Capacity Costing**

One way to implement the commandments of capacity costing would be to create a compromise between variable and absorption costing: immediately expense the costs of capacity not demanded by customers, including strategic, contractual, and excess capacity. However, a far more common method is to set the activity base equal to the firm’s practical capacity to produce output for customers, rather
The Death Spiral and Capacity Costing

than the level of expected activity. **Practical capacity** is the amount of total driver use, assuming competent management but recognizing that a number of unavoidable factors can limit capacity, including the need to maintain surge and buffer capacity. For example, consider a firm that incurs $2,400,000 in fixed costs to maintain the capacity to produce 100,000 units per year under ideal conditions. However, because the firm wants to promise rapid delivery, they maintain idle surge and buffer capacity of 10,000 units each. Recent economic conditions have reduced their actual levels of production to only 50,000 units. Their capacity breakdown would therefore be:

| Total Capacity | 100,000 |
| Surge Capacity | <10,000> |
| Buffer Capacity | <10,000> |
| **Practical Capacity** | <80,000> |
| Projected Production | 50,000 |
| Excess Capacity | 30,000 |

Intelligent capacity costing would dictate that they use 80,000 units of practical capacity as their activity base, rather than the 50,000 units of projected production (utilized capacity). This choice results in a predetermined fixed overhead rate of $30/unit ($2,400,000/80,000 units), and has several useful effects:

1. Customers bear the cost of the surge and buffer capacity they demand (since they demand rapid delivery), because those components of idle capacity reduce the activity base, and therefore increase the overhead applied to each unit produced. Using total capacity as an activity base would reduce the predetermined overhead rate to $24/unit ($2,400,000/100,000 units).
2. Customers do not bear the cost of excess capacity because that component of idle capacity does not reduce the activity base—losing more customers would not increase overhead applied to the remaining units sold. Excluding the 30,000 units of excess capacity from the activity base by using projected production would boost the predetermined overhead rate to $48/unit ($2,400,000/50,000 units).
3. The firm avoids a Death Spiral by divorcing the activity base from the projected level of production!

Note that implementing capacity costing by using practical capacity as the activity base can be counterintuitive: we charge customers for idle surge and buffer capacity by excluding it from the activity base, and avoid charging customers for excess capacity by including it in the activity base! Remember, the smaller the activity base, the higher the reported cost per unit!
Nested Capacity, the Cost Hierarchy, and Activity-Based Costing

Nested Capacity

The Cost-Volume-Profit (CVP) model assumes that every cost is either fixed or variable. In fact, most businesses include costs that are neither fixed nor variable, but rather combine aspects of each.

Consider Joe, who runs a lawn-mowing service. Joe must first invest in a lawn mower and a truck to transport it, at cost of $F_{equip}$. Then he must use gas and travel time to get to the location of each client, at a cost of $B_{client}$; for simplicity we will assume these costs are roughly the same for each address. Once he has traveled to the right address, he can mow $Q$ square meters (SQM) of lawn at a variable cost of $v$ per square meter. The total cost function for this simple business would be:

$$\text{Cost} = F_{equip} + B_{client} \cdot \text{Clients} + vQ$$

Is the cost of traveling to a client fixed or variable? Like a fixed cost, this cost doesn’t change as a function of how many square meters he mows—the cost of traveling to a client is no higher if the lawn is larger. Like a variable cost, it does vary with a measure of his activity—the number of different lawns he mows.

The terms “fixed” and “variable” don’t quite work for the lawn-mowing business because it requires nested capacity. The purchase of a lawn mower and truck provides the capacity to travel to clients with the appropriate equipment; traveling to a client provides the capacity to mow square meters of the client’s lawn. This nesting of capacity implies that Joe’s business has three different levels of cost. The equipment costs are the highest level, and the per-SQM cost is the lowest level; these are the familiar fixed and variable costs captured by CVP. Traveling to the client is an intermediate-level cost, which is fixed when viewed from the perspective of the lower level (costs of traveling to a client do not vary with the number of square meters) and variable when viewed from the perspective of the higher level (travel costs increase with the number of clients). Here is a figure illustrating the business’s capacity nesting:
Intermediate-level costs can distort reported margins by forcing high-volume business lines to cross-subsidize low-volume lines. To see how, assume that the firm allocates overhead on the basis of square meters. The firm would then include the costs of both equipment and travel into an overhead pool, and allocate the same amount of those costs to each square meter mowed. This simple system allocates the same overhead to one 1,000 SQM job as it does to ten 100 SQM jobs. However, the true cost of the single large job is substantially lower than that of the ten smaller jobs, because it requires only one trip to the client’s location.

**Identifying Nested Capacity and Cost Levels**

You can identify nested capacity in your business by constructing sentences that follow this basic format:

> We need to [engage in costly activity X] in order to [repeat costly activity Y], but much of the cost of [X] is largely independent of [how many repetitions of Y we undertake].

If such a sentence describes your production process, X is a higher-level activity than Y. Settings with many activities that can be nested (X enables Y, which enables Z, which enables W) are said to have a cost hierarchy.

Before discussing cost hierarchies for entire businesses, let’s look at individuals who engage in two common time-consuming activities: air travel and public speaking. For these examples, the “cost” is measured entirely in terms of time.

**Air Travel**

Melissa travels almost every week, and does everything she can to minimize travel time. She also knows that travel time is not simply a matter of how many miles she is flying. A short trip split into many legs might take far longer than a nonstop flight across the continent. Measuring cost purely in terms of the minutes it takes Melissa to get from home to hotel, she can identify a very natural hierarchy: trip-level costs at the top, leg-level costs in the middle, and mile-level costs at the bottom.
Nested Capacity, the Cost Hierarchy, and Activity-Based Costing

Trip-level costs are the time it takes to make it possible to fly one or more legs of an itinerary: driving to an airport, checking in, getting through security, retrieving luggage at the destination, and driving to her hotel. Melissa needs to perform these activities to fly any number legs of her trip, and the cost (time) devoted to trip-level activities does not vary with how many legs she flies.

Leg-level costs capture the time it takes to make it possible to fly a single mile on a flight leg: waiting in the airport before the flight, boarding the plane, waiting on the tarmac for the plane to take off, waiting on the tarmac for the plane to taxi to the gate, disembarking, and waiting in the airport. Melissa needs to engage in these activities to fly any number of miles on a leg, and the cost (time) devoted to leg-level activities does not vary with how many miles are in the leg.

Mile-level costs capture the time it takes to fly a single mile, once the time associated with leg-level and trip-level activities has already been incurred.

Public Speaking

John works in human resources at a large corporation. He must frequently give presentations to his colleagues to promote new initiatives, policies, and best practices. He is well aware that the time he devotes to this part of his job goes far beyond the time he spends actually speaking in public. He also sees that the time demands fall naturally into a hierarchy: expertise-level costs at the top, speech-level costs in the middle, and delivery-level costs at the bottom.

Expertise-level costs include the time John must devote to becoming an expert in the general subject areas he will speak on. Developing that expertise takes about the same amount of time, regardless of the number of speeches he will be giving.

Speech-level costs include the time John must devote to crafting a particular speech for a particular audience. Developing a speech takes about the same amount of time, regardless of how many times he will be delivering it.

Delivery-level costs include the time it actually takes John to deliver a speech to an audience. This might include travel time, preparation time, and the duration of the actual speech itself.

Activity-Based Costing and the “Textbook” Hierarchy

Activity-Based Costing (ABC) is a costing system designed to accommodate cost levels that arise from nested capacity. ABC’s ability to account intelligently for intermediate-level costs has made it one of the most popular innovations in cost accounting since the early twentieth century. Every business process has a slightly different cost hierarchy, but most discussions begin with a simple “textbook” hierarchy of
four levels: facility level, product level, batch level, and unit level. We can identify these levels in a manufacturing company as follows:

- **Facility Level.** We need to own and operate our own facilities to run a business that can support one or more product lines; within a broad relevant range, the costs of the facility are largely independent of how many different product lines we sell.

- **Product Level.** We need to develop engineering specifications, find suppliers and producers, and keep records for each of our product lines; the costs of supporting a product line are largely independent of how many production runs we initiate.

- **Batch Level.** Every time we produce some units in a particular product line, we must set up our machines appropriately, calibrate them to the specifications the line requires, clean residue from prior production runs, and maintain appropriate paperwork to schedule the production run and track the transfer of goods into and out of the production floor. The costs of initiating a production run are largely independent of how many units the run produces.

- **Unit level.** Every time we make a unit as part of a production run, we incur additional costs for labor, materials, and shared resources.

Here is a figure illustrating the textbook hierarchy of costs:

![Cost Hierarchy Diagram]

**Activity-Based Costing vs. Volume-Based Costing**

The basic mechanisms of ABC systems resemble those of other overhead allocation systems. We collect total overhead costs into a variety of cost pools, each associated with a single driver, called an activity trigger (the event that is said to trigger all of the activities whose costs are captured in the cost pool). We divide the projected overhead costs in the pool by the trigger’s activity base to calculate a predetermined overhead rate for the pool, and use that rate to apply the pool’s costs to cost objects.

Consider, for example, a professional services firm that allocates $1,000,000 of overhead across an activity base of 20,000 billable hours (BH). This simple system could be depicted as follows:
Simple Overhead Allocation

Now imagine the firm adopts an ABC system by creating four overhead pools, as follows:

- **Facility costs.** This pool includes the costs of maintaining and heating the office buildings, and the cost of satisfying professional certification requirements. Pool costs of $300,000 are allocated on the basis of 20,000 BH, yielding a predetermined overhead rate of $15/BH.

- **Service line costs.** This pool includes the costs of keeping the firm at the forefront of expertise required for each different type of professional service provided, such as training, software, licenses, subscriptions, and travel to conferences. Pool costs of $200,000 are allocated equally to each of the firm’s 10 service lines, yielding a predetermined overhead rate of $20,000 per service line.

- **Client costs.** This pool includes the costs associated with attracting and maintaining clients. Such costs include travel and entertainment expenses, billing, maintaining and updating mailing lists, and keeping records for tax and other purposes. Pool costs of $150,000 are allocated equally to each of the firm’s 150 clients, yielding a predetermined overhead rate of $1,000 per client.

- **Unit costs.** This pool includes the costs associated with each incremental hour of professional services provided, including secretarial support, duplicating, travel to clients, phone service, and software. Pool costs of $350,000 are allocated on the basis of 20,000 BH, yielding a predetermined overhead rate of $17.50/BH.
The following figure depicts the new cost system:

This figure is very similar to those depicting traditional multipool systems. But rest assured that Activity-Based Costing (ABC) is not simply a fancier name for the same costing systems we studied before. Those systems effectively revealed hidden drivers and hidden differences, but the additional drivers and differences were all unit-level costs. As a result, doubling the volume of hours in a job with a constant mix of partner and associate hours would double the cost. Systems that use only unit-level costs are typically called volume-based costing systems, or traditional costing systems, to distinguish them from ABC systems.

Because they use high-level cost drivers, reported costs in ABC systems typically rise much more slowly than the volume of production. This is true because the inputs demanded by the contract itself determine only part of the cost of an individual contract. Using the overhead rates above, each BH will result in the application of $15 of Facility costs and $17.50 in Hours costs. The cost of a job with 100 hours would include $15 \times 100 + $17.50 \times 100 = $3,250 as a result of the overhead based on BH. However, the rest of the cost would depend on the nature of other jobs in the firm. For example, assume that the contract is one of two for a particular client, and one of ten in a particular service line. The job would then get charged an additional $1,000/2 = $500 for its client cost, and an additional $20,000/10 = $2,000, for a total cost of $3,250 + $500 + $2,000 = $5,750. This job therefore generates an average of $57.50 in overhead for each BH.

Why does this particular job generate an overhead rate higher than the average job ($57.50/BH vs. $50/BH under the original system)? Because it has relatively few hours, but still must bear half the cost associated with a single client, and one-tenth the cost of a single service line. Other jobs will be less
expensive per BH—those with many BHs, those with clients who have many other contracts with the firm, and those in very popular service lines. Such jobs can share the higher-level service line and client costs with many other contracts, and spread those costs over many service hours.

**The Danger of Low-Volume Business**

One reason ABC systems are popular is their ability to help businesses recognize the high true costs and low true margins of low-volume business. Consider a simple business, Hypertrophy, Inc., which makes trophies for 15 different types of competitions. As shown in the spreadsheet in the Appendix, Hypertrophy can sell 1,000 trophies a month for popular sports such as volleyball, basketball, and hockey; they sell only 50 trophies a month for chess, curling, and tiddlywinks. They sell trophies for other sports in quantities of 500, 200, 150, or 100 per month. The direct cost of making each trophy is $10, regardless of the type of trophy. However, the firm devotes considerable resources per month to maintaining each product line through a combination of design, engineering, and supply-chain activities. These product-level costs total $15,000 a year. The company can produce trophies in batches of 100. The cost of scheduling, setting up before a batch, and cleaning up after is about $200. The company makes a total of 58 batches per month, for a total batch-level cost of $11,600.

Assume Hypertrophy allocates the same overhead to each trophy. The product-level and batch-level costs total $26,600, which when spread over a total volume of 5,600 units generates an average overhead cost of $4.75 per trophy; adding in the direct cost of $10, the average trophy costs $14.75.

However, this analysis ignores the fact that the low-volume trophies generate just as much overhead per product line and per batch as the high-volume trophies. From an ABC perspective, 1,000 basketball trophies drive far less overhead than $4.75 per unit. An ABC system would allocate this high-volume product line $1,000 in product-level costs and $2,000 in batch-level costs (since it takes 10 batches to make 1,000 units), for a total of $3,000 in overhead for the entire line, or $3.00 per trophy.

In contrast, an ABC system would allocate the low-volume curling trophies the same $1,000 in product-level costs and a full $200 in batch-level costs, because producing only 50 trophies still entails setting up for a production run. The $1,200 is spread over only 50 units, for a whopping $24 of overhead per trophy.

Now assume that the firm, which is not implementing ABC, charges a 10% markup over the reported cost of $14.75 for each trophy, resulting in a price of roughly $16.25 per trophy. The high-volume products are generating a “true” average margin of $16.25 − $13.00 = 3.25 per unit, or $3,250 for the entire volume of 1,000 units. The low-volume products are generating an average margin of about $16.25 − $34.00 = −17.75 per unit, or −$887.50 for the entire volume of 50 units. Producing curling trophies is costing $887.50 in profit per month!

We can use our knowledge of the true ABC costs to create a graph of the cumulative profit margin for each line of business, with each segment of the line indicating the additional profit or loss generated by a single product line. The total profit of the firm is $8,000, which is consistent with the overall 10%
markup on its $81,900 in cost. However, they could earn far more money by dropping their low-volume product lines. Assuming the firm could actually shed $1,000 of resource spending for each product line eliminated, and shed $200 for each production run it didn’t initiate, the firm would be able to earn over $13,000 per month by dropping all but their six most popular product lines.
## Nested Capacity, the Cost Hierarchy, and Activity-Based Costing

<table>
<thead>
<tr>
<th>Products</th>
<th>Units/Batch</th>
<th>Units</th>
<th>Batches</th>
<th>Total True Cost (ABC)</th>
<th>True Cost/Unit</th>
<th>Volume-Based (Single Driver) Cost</th>
<th>Reported Margin</th>
<th>Actual Margin</th>
<th>Actual Margin/Unit</th>
<th>Cumulative Volume</th>
<th>Cumulative Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basketball</td>
<td>1</td>
<td>100</td>
<td>1000</td>
<td>$13,000.00</td>
<td>$13.00</td>
<td>$14,750.00</td>
<td>$1,475.00</td>
<td>$3,225.00</td>
<td>$3.23</td>
<td>$1,000.00</td>
<td>$3,225.00</td>
</tr>
<tr>
<td>Volleyball</td>
<td>1</td>
<td>100</td>
<td>1000</td>
<td>$13,000.00</td>
<td>$13.00</td>
<td>$14,750.00</td>
<td>$1,475.00</td>
<td>$3,225.00</td>
<td>$3.23</td>
<td>$2,000.00</td>
<td>$6,450.00</td>
</tr>
<tr>
<td>Baseball</td>
<td>1</td>
<td>100</td>
<td>1000</td>
<td>$13,000.00</td>
<td>$13.00</td>
<td>$14,750.00</td>
<td>$1,475.00</td>
<td>$3,225.00</td>
<td>$3.23</td>
<td>$3,000.00</td>
<td>$9,675.00</td>
</tr>
<tr>
<td>Football</td>
<td>1</td>
<td>100</td>
<td>1000</td>
<td>$7,000.00</td>
<td>$14.00</td>
<td>$7,375.00</td>
<td>$737.50</td>
<td>$1,112.50</td>
<td>$2.23</td>
<td>$3,500.00</td>
<td>$10,787.50</td>
</tr>
<tr>
<td>Cycling</td>
<td>1</td>
<td>100</td>
<td>500</td>
<td>$7,000.00</td>
<td>$14.00</td>
<td>$7,375.00</td>
<td>$737.50</td>
<td>$1,112.50</td>
<td>$2.23</td>
<td>$4,000.00</td>
<td>$11,900.00</td>
</tr>
<tr>
<td>Golf</td>
<td>1</td>
<td>100</td>
<td>500</td>
<td>$7,000.00</td>
<td>$14.00</td>
<td>$7,375.00</td>
<td>$737.50</td>
<td>$1,112.50</td>
<td>$2.23</td>
<td>$4,500.00</td>
<td>$13,012.50</td>
</tr>
<tr>
<td>Science Fair</td>
<td>1</td>
<td>100</td>
<td>200</td>
<td>$3,400.00</td>
<td>$17.00</td>
<td>$2,950.00</td>
<td>$295.00</td>
<td>-$155.00</td>
<td>-$0.78</td>
<td>$4,700.00</td>
<td>$12,857.50</td>
</tr>
<tr>
<td>Tennis</td>
<td>1</td>
<td>100</td>
<td>200</td>
<td>$3,400.00</td>
<td>$17.00</td>
<td>$2,950.00</td>
<td>$295.00</td>
<td>-$155.00</td>
<td>-$0.78</td>
<td>$4,900.00</td>
<td>$12,702.50</td>
</tr>
<tr>
<td>Marathons</td>
<td>1</td>
<td>100</td>
<td>200</td>
<td>$3,400.00</td>
<td>$17.00</td>
<td>$2,950.00</td>
<td>$295.00</td>
<td>-$155.00</td>
<td>-$0.78</td>
<td>$5,100.00</td>
<td>$12,547.50</td>
</tr>
<tr>
<td>Track</td>
<td>1</td>
<td>100</td>
<td>150</td>
<td>$2,900.00</td>
<td>$19.33</td>
<td>$2,212.50</td>
<td>$221.25</td>
<td>-$466.25</td>
<td>-$3.11</td>
<td>$5,250.00</td>
<td>$12,081.25</td>
</tr>
<tr>
<td>Skeet Shooting</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>$2,200.00</td>
<td>$22.00</td>
<td>$1,475.00</td>
<td>$147.50</td>
<td>-$577.50</td>
<td>-$5.78</td>
<td>$5,350.00</td>
<td>$11,503.75</td>
</tr>
<tr>
<td>Frisbee</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>$2,200.00</td>
<td>$22.00</td>
<td>$1,475.00</td>
<td>$147.50</td>
<td>-$577.50</td>
<td>-$5.78</td>
<td>$5,450.00</td>
<td>$10,926.25</td>
</tr>
<tr>
<td>Curling</td>
<td>1</td>
<td>100</td>
<td>50</td>
<td>$1,700.00</td>
<td>$34.00</td>
<td>$737.50</td>
<td>$73.75</td>
<td>-$888.75</td>
<td>-$17.78</td>
<td>$5,500.00</td>
<td>$10,037.50</td>
</tr>
<tr>
<td>Chess</td>
<td>1</td>
<td>100</td>
<td>50</td>
<td>$1,700.00</td>
<td>$34.00</td>
<td>$737.50</td>
<td>$73.75</td>
<td>-$888.75</td>
<td>-$17.78</td>
<td>$5,550.00</td>
<td>$9,148.75</td>
</tr>
<tr>
<td>Tiddlywinks</td>
<td>1</td>
<td>100</td>
<td>50</td>
<td>$1,700.00</td>
<td>$34.00</td>
<td>$737.50</td>
<td>$73.75</td>
<td>-$888.75</td>
<td>-$17.78</td>
<td>$5,600.00</td>
<td>$8,260.00</td>
</tr>
</tbody>
</table>
Activity-Based Pricing and Supply-Chain Surplus

Introduction

ABC’s focus on activities makes it useful for much more than calculating costs. Most companies with ABC systems use them as a basis for Activity-Based Management (ABM) and Activity-Based Pricing (ABP). The relationship between ABC and ABM is often represented by a figure such as this one (a simplified version of Exhibit 5-11 in Hilton’s Managerial Accounting textbook):

Activity-Based Costing & Management

Rather than simply using ABC to assign resource costs to cost objects (the vertical arrows), ABM uses the cost information provided by an ABC system to reduce the frequency of events that trigger activities, and to reduce the cost of those activities (the horizontal arrows). ABM and ABP systems are excellent illustrations of the unifying principle in these essays: there IS a free lunch, if you have a reporting system that allows you to sniff it out. You can use ABM systems to eliminate activities that don’t pass a cost-benefit analysis, and use the savings to buy a free lunch. You can use ABP to raise prices on products and services that demand many higher-level (product or batch) costs, and lower prices on those that don’t. These new prices can increase the total value your organization and its customers creates; the two of you can pool that added wealth and use it buy each other lunch!
Activity-Based Management

Consider a firm that makes 10,000 widgets a day, in separate production runs of 1,000. Each production run triggers $2,000 of labor costs to clean and recalibrate machines and record production and accounting information. The firm might be able to reduce these batch-level costs in two ways: by reducing the time and effort associated with each individual setup, and by reducing the number of production runs needed to produce 10,000 widgets.

Activities that can be reduced in cost or eliminated without affecting customers’ experience are often referred to as non-value-added activities. For example, assume that firms do not care about the size or frequency of production runs. The firm can then reduce batch-level costs by reengineering the production process (or simply rescheduling it) to increase each production run from 1,000 to 2,000 units. Alternatively, the firm might be able to reduce the cost of setting up for each production run, reducing the cost from $2,000 to $1,500. If these changes have no negative effects on customers’ experience, the firm should try to make them.

Activity-Based Pricing and Insufficient-Value-Added Activities

Firms are often hard-pressed to find many activities that add no value to customers, especially after they have implemented their ABC-ABM system and picked the low-hanging fruit. It is much more common for firms to engage in activities that add some value to some customers, but not enough to justify their costs. Firms can identify these insufficient-value-added activities by building activity costs into the prices they pay (Activity-Based Pricing). Customers may value frequent delivery, but are they willing to pay the true cost of the extra production runs they demand? Customers may value being able to purchase a low-volume product that meets their needs better than similar higher-volume products, but are they willing to pay the true costs of sustaining a product line that are spread over a very small customer base? If not, these activities provide insufficient value to customers, and they will fade away as demand for them falls in response to high prices.

Supply Chain Surplus and Integrative Bargaining

ABC doesn’t just alter prices; it alters the nature of bargaining between customers and suppliers along the supply chains. The collective goal of firms in a supply chain is to maximize the supply chain surplus, defined as the value the supply chain’s output provides to end consumers, less the aggregate costs of the producers in the chain. The goal of an individual firm in the supply chain is twofold: to identify production choices for itself and its immediate suppliers and customers that increase the total supply chain surplus (“making the pie bigger”), and to ensure that the firm itself maximizes its own share of that surplus (“getting a bigger share of the pie”). Firms that focus exclusively on getting a bigger share of the pie are engaged in distributive bargaining (i.e., bargaining over how the distribution will be distributive), while those that also focus on making the pie bigger are engaged in integrative bargaining (i.e., bargaining in a cooperative way to make the pie larger for everyone).
Activity-Based Pricing and Supply-Chain Surplus

By allowing firms to identify and eliminate activities that add insufficient value, ABP allows them to engage in integrative bargaining. This results not only in greater supply-chain surplus, it allows firms to establish more cooperative relationships with their suppliers and customers. In purely distributive bargaining, one side’s gain is another side’s loss. As a result, it rarely makes sense for firms engaged in distributive bargaining to share information. Information sharing can be extremely valuable in integrative bargaining, however, because it allows firms to identify how they might alter their behaviors in a way that increases the size of the pie (the supply-chain surplus).
Part VII: Coordination
Credit and Blame, Authority, Responsibility, and Accountability

One of the most contentious issues in any organization is how to assign credit and blame to employees. Credit and blame are conclusions about how individual actions caused good or bad observed outcomes. Such conclusions are difficult to draw even when a single person works independently, because observed outcomes reflect a combination of individual actions, circumstances beyond the person’s control, measurement error. Assigning credit and blame is even more difficult when many people are working together to achieve a common purpose, making it unclear whose actions caused which effect.

Reporting systems can help organizations assign credit and blame, but successful organizations recognize that as conclusions, such assignments represent the last step in the performance evaluation process—a step that can be made much easier by careful attention to the two preceding steps: assigning authority and assigning responsibility.

Authority is the right to make and implement decisions, and is the first step in designing any organization. Clearly assigning authority makes assigning credit and blame in a large organization more similar to assigning credit and blame to an independent worker (easier, but still not easy).

The next step is to assign responsibility, which is a set of duties to:

- **Know** how well your portion of the business is performing
- **Explain** deviations from expected performance to others when appropriate
- **Propose** responses expected to improve performance, in light of your explanations

Clearly assigning responsibility allows straightforward tests of how well managers performed the duties that responsibility entails, even if it is hard to know whether failure to fulfill a duty caused bad outcomes. Note that managers who are assigned authority are also assigned responsibility, because no one can make intelligent decisions without knowing current performance and being able to explain it and propose responses to it. Managers may have responsibility without authority, if their immediate superiors retain the right to make decisions, relying on their subordinates to keep them informed.

Once an organization has assigned responsibility and authority, it can ask who should be rewarded or punished based on each measure. This allocation of reward and punishment is called accountability. Accountability can be either explicit or implicit. Accountability is explicit if the rewards and punishments
are tied to predetermined measures of performance (as in many pay-for-performance plans). Accountability is implicit if those who decide rewards and punishments make no commitment as to how they will do so, typically determining them through a system of performance reviews and subjective evaluations. Whether explicit or implicit, accountability ideally tracks credit and blame: a manager who has caused good performance should be rewarded through bonuses, raises, promotions, or commendations; a manager who has caused bad performance should be punished through low pay, reprimands, or even termination. However, the difficulty of causal attributions means that accountability rarely tracks credit or blame as closely as we might hope.

Responsibility Centers

One common and effective way to assign authority, responsibility, and accountability is to subdivide an organization into responsibility centers, semi-autonomous organizational units, each of which is (1) directed by a single party (sometimes a committee) with authority for its performance, and (2) evaluated based on its own set of managerial reports. The directors of the center are responsible for the complete set of financial statements, but they may further isolate responsibility by identifying one subordinate to be responsible for each line item on the center’s performance report.

### Best Practices in Isolating Responsibility

Isolate responsibility by making a single person responsible for each line item in a managerial report. This manager is said to “own” the line item, even if his or her superior is the one with the authority to make decisions.

Owning a line item does not mean poor performance on that item is “your fault” (an attribution of credit or blame). Processing costs may be too high because oil prices rose in response to geopolitical tensions. Clearly this is not your fault. You are living up to your responsibility if you inform your superiors of a problem, explain the cause, and propose that the firm buy oil price futures, raise selling prices, or shift to less oil-intensive means of production.

### Types of Responsibility Centers

Managerial accountants typically describe responsibility centers according to the elements of financial performance its directors are responsible for. Here are the most common types of responsibility centers:

- **Investment centers** are evaluated in practically the same way as stand-alone businesses, being held responsible for their investment base as well as their profit. Investment centers report both an income statement and a balance sheet, and directors are held responsible for Return on Assets and Return on Investment, as well as Revenue, Expense, and Profit.
- **Profit centers** are evaluated on the basis of reported profit, but directors are not held responsible for the assets under their control. Profit centers therefore must generate a
Complete income statement, but may not generate a balance sheet. Thus, directors are responsible for Revenue, Expense, and Profit, but not Return on Assets or Return on Investment.

- **Cost centers** are evaluated on the basis of reported production costs, but are not responsible for revenue or investment base. Cost centers must therefore generate only a partial income statement that reports Cost of Goods Sold and other expenses, but not Revenue or a Balance Sheet. Many production units and support units (such as Technology Services) are designed as cost centers, because they transfer their output to other centers without recognizing any revenue.

- **Revenue centers** are evaluated on the basis of reported revenue, but are not responsible for their production costs or investment base. Revenue centers must therefore generate only a partial income statement that reports Revenues, but not Cost of Goods Sold or a Balance Sheet. Many sales units are designed as revenue centers, because they are not charged for the goods and services they receive from other centers.

Investment centers are better at isolating responsibility than the other types of centers are, because their directors are least able to argue that the actions of other investment centers influenced their center’s performance. For example, the director of a revenue center can claim that sales are low because a cost center provided low-quality goods to sell. The director of a cost center can argue that their costs are high because they gave a revenue center high-quality goods to sell. The director of a profit center can argue that profits were low because they were not given adequate resources to run their business. The director of an investment center cannot easily make these arguments, because the center is responsible for its own revenue, costs, and investment base.

**An Example**

* Totally Nuts creates three products through the course of three production departments, as shown in the figure below. The primary raw materials (nuts) are first processed in a Receiving Department, which cleans them and removes those that don’t meet quality standards. They transfer the processed nuts to Roasting, where they are coated with oil, salt, and spices and heated in large ovens. The roasted nuts are transferred to the Coating Department, which applies candy or chocolate coatings and sells the product to candy stores.
If it were a stand-alone company, Totally Nuts could be viewed as an investment center, responsible for the costs incurred by each department, the revenue received by Coating, and the investment in assets used for processing. However, it would be easier to evaluate each element of production if Totally Nuts viewed each department as a responsibility center. Assuming that top management retains authority over investment decisions, Receiving and Roasting would both be cost centers, because they incur costs but have no revenues. Coating could be treated as a profit center, because it receives revenue and incurs costs.

But this treatment of Totally Nuts’ departments is unsatisfactory in several ways. First, it would be preferable to treat Receiving and Roasting as investment or profit centers, rather than cost centers, to clarify the division of responsibility. As it stands, each department can reduce its costs by transferring lower-quality goods to the next department. The firm treats Coating as a profit center, but it has no authority (and should have no responsibility) for most of the costs of creating coated nuts, which are incurred in other departments.

Constructing responsibility centers becomes more difficult when you allow the directors of each department more authority. Assume, for example, that each department can sell its goods directly to outside customers, and each department can buy goods from outside suppliers, not just from the preceding department, as shown in the following figure.
Responsibility Centers, Transfer Pricing, and Decentralization

### Totally Nuts

Now Receiving and Coating can have some revenue, but treating them as profit centers will encourage them to sell to outside customers (for which they get revenue) rather than transferring their output to the next department (which may be preferable for the firm, but does not result in revenue to the department). Similarly, the departments have no incentive to purchase goods from outside suppliers, since that causes them to incur a cost, whereas accepting a transfer from another department does not. Fortunately, organizations have a tool that helps them solve this problem: the transfer price.

#### Transfer Prices and the Internal Supply Chain

Totally Nuts can isolate responsibility centers more effectively by setting transfer prices for goods and services that are transferred from one center to another. Transfer prices allow the firm to view their internal responsibility centers as forming links in an internal supply chain. An internal supply chain differs from an external supply chain in that, by definition, an internal one involves transfers of goods and services between two related parties, who each serve as an agent for the same principals—those who own or direct the organization.

Transfer prices can convert cost and revenue centers into profit or investment centers. For example, Totally Nuts could set a price for the transfer of processed nuts from Receiving to Roasting, and another price for the transfer of roasted nuts from Roasting to Coating. Then Receiving, Roasting, and Coating all become profit centers, recognizing the costs of goods transferred in (or bought from outside suppliers) and recognizing the revenue of goods transferred out (or sold to outside customers).

Adding a transfer price for the transfer of capital assets from headquarters to each department (such as the factory and equipment) allows each department to be viewed as an investment center; each
recognizes the interest expense for the capital used to support its operations, and reports the value of
the capital assets on each center’s Balance Sheet.

Purposes and Types of Transfer Prices

Firms have many options for setting transfer prices; the best choice depends on the firm’s goals. At the
simplest level, firms can use transfer prices to track the flow of costs. However, they can also use them
to hold cost centers responsible for production efficiencies, to improve sourcing decisions (whether the
firm should make or buy key inputs), and to capture the value each center creates. The following pages
describe the transfer prices most appropriate for achieving these goals, in order of increasing
sophistication.

1. Normal Cost Transfer Prices Help Track the Flow of Costs and Responsibility

One simple and common transfer pricing method is to set the transfer price equal to the cost of the
good or service already captured by the transferring center under normal absorption costing. These
normal (absorption) cost transfer prices (usually called normal cost transfer prices or normal full-
cost transfer prices) serve two purposes. First, they ensure that the cost of the product or service
ultimately provided to external parties captures the cost of the transferred product or service. In Totally
Nuts, the Receiving department treats its sorted nuts as a cost object, to which it applies the costs of
raw materials and processing. When it transfers raw nuts to Coating, the firm will credit the finished
goods inventory of the receiving department, just as if it sold the nuts to an outside buyer. The
difference lies in the account debited; for a transfer, Totally Nuts will debit the raw materials account for
Roasting. Assuming that it cost $120 to process a batch of nuts, the entry for the transfer is:

\[
\begin{align*}
\text{Dr.} & \quad \text{Raw Materials Inventory—Roasting} & \quad \$120 \\
\text{Cr.} & \quad \text{Finished Goods Inventory—Receiving} & \quad \$120 \\
\end{align*}
\]

To record the transfer of a batch of nuts from Receiving to Roasting

Normal costing transfer prices serve another purpose as well. Transfers of goods almost always entail a
transfer of responsibility, too. For example, when Totally Nuts’s Receiving department transfers raw
nuts to Roasting, it also transfers the duty to protect the nuts from spoilage and theft. Recording this
transfer of responsibility is essential to the organization’s control system, but it cannot be entered into a
double-entry bookkeeping system without specifying a dollar amount.

2. Standard Cost Transfer Prices Help Isolate Responsibility for Production Efficiencies

Many firms use transfer prices to isolate responsibility for production efficiencies. Normal cost transfer
prices do a poor job of isolating this responsibility, for the simple reason that the transferring center
applies all of its costs to the inventory it transfers to the next center. For example, assume that
Receiving was quite inefficient in sorting nuts, causing needlessly high labor costs per kilogram of nuts.
Inefficiency would be unlikely to increase the revenue received in a transaction with an outside buyer,
so the increase in Cost of Goods Sold recorded upon sale would result in a lower reported margin.
Responsibility Centers, Transfer Pricing, and Decentralization

However, if the nuts were transferred to Roasting before their sale, Receiving would simply pass the higher costs on to Roasting. Roasting would then report a lower margin because they would include higher raw materials costs in their Cost of Goods Sold. As the last department in the internal supply chain, Coating’s margin will reflect the production efficiencies of all of the preceding departments. Because no department other than Receiving has sole (isolated) responsibility for its reported margins, they lack a strong incentive to reduce their costs.

Organizations can isolate cost responsibility more effectively by setting a transfer price equal to the standard cost that reflects the amount each center of output should have cost for its phase of production. Assume, for example, that it should have cost Receiving $100 for every batch of nuts they transfer to Roasting, according to the production standard set by Totally Nuts, and that the $20 of extra cost is entirely due to inefficient use of labor. Under standard cost transfer pricing, Receiving would transfer only $100 into Roasting, with the $20 difference recorded as a labor efficiency variance, which is treated as an expense or a loss. This standard costing method can be used for transactions as well, separating the cost of Receiving’s inefficiency from Cost of Goods Sold, which reflects what the sold output should have cost, rather than what it actually did cost. The entries for the transfer are as follows.

Dr. Raw Materials Inventory—Roasting $100
Dr. Labor Efficiency Variance $20
Cr. Finished Goods Inventory—Receiving $120

To record the transfer of a batch of nuts from Receiving to Roasting and the recognition of the related labor efficiency variance

Setting and recording standard costs and variances is a complicated topic, and I discuss it in more detail in a later essay.

3a. Variable Cost Transfer Pricing Improves Sourcing Decisions When Centers Have Excess Capacity

Many firms use transfer prices to allow managers to make better decisions about whether to purchase goods from an outside supplier, rather than from another center within the firm. The owner’s of Totally Nuts, for example, might allow the Roasting manager to purchase raw nuts from outside processors. The appropriate make-or-buy decision depends on whether the total costs of Totally Nuts would be higher if Receiving processed the nuts or if Roasting made a cash payment to another firm. Because processing an additional batch of nuts does not cause Receiving to incur additional fixed costs, full (absorption) cost transfer prices would mislead the Roasting manager into buying processed nuts rather than requesting them from Receiving. As long as Receiving has the capacity to process an additional batch, a variable cost transfer price produces better make-or-buy decisions, because it forces the Roasting manager to bear the firm’s true incremental cost of asking Receiving to produce an additional batch.

3b. Market Price Transfer Pricing Improves Sourcing Decisions When Centers Lack Excess Capacity

If Receiving lacks the extra capacity to process one additional batch of nuts, the incremental cost of requesting a batch is not the variable cost (since the total number of batches made does not change),
but the loss of a sale that would have been made to an outside party. Firms can capture this opportunity cost by setting a market-based transfer price that equals the revenue lost by transferring the batch, net of any marginal costs related to the sale.

4. **Cost-Plus and Negotiated Transfer Prices Help Isolate Responsibility for Value Creation**

Many organizations use transfer prices to isolate each center’s contribution to the total value created by their transactions with outside suppliers. Just as market prices separate an entire chain’s surplus into margins earned by each link, transfer prices can do the same for internal supply chains, as long as they allow each center to transfer goods above reported cost (as long as they are contributing value to the ultimate product or service). Assuming that Receiving transfers a batch of nuts for $150 when it cost them only $120, the appropriate entries would be as follows:

- **Dr.** Cost of Goods Sold—Receiving $120  
  **Cr.** Finished Goods Inventory—Receiving $120
  *To record Receiving’s expense for “selling” goods to Roasting*

- **Dr.** Transferred Payments from Roasting—Receiving $150  
  **Cr.** Revenue from Roasting—Receiving $150
  *To record Receiving’s revenue and consideration for “selling” goods to Roasting*

- **Dr.** Raw Materials Inventory—Roasting $150  
  **Cr.** Transferred Payments from Roasting—Roasting $150
  *To record Roasting’s cost of “purchasing” goods from Roasting*

Note that the transfer price involves Roasting paying a “consideration” to Receiving in the form of an account receivable for Receiving and an account payable for Roasting. Unlike a true transaction, Roasting never parts with a dollar of true resources; this consideration is simply a means of allowing Receiving to report revenue, and therefore to be treated as a profit center.

The simplest method of isolating responsibility for value creation is to use **cost-plus transfer prices** that add a percentage markup to a cost reported under any of the methods described above (most commonly normal or standard absorption costs). While the margin can be determined by an executive to whom the involved parties both report, many firms hand that power to the involved parties themselves. Because all parties are evaluated on the basis of reported margins, they have enough of an adversarial relationship to allow negotiations over the transfer price to approximate the outcome of an arm’s-length transaction.

**Transfer Prices, Knowledge, and Decentralization**

As transfer pricing becomes more sophisticated, it becomes a more effective tool for allowing the directors of each center to make their own decisions (decentralization), as opposed to having all
Responsibility Centers, Transfer Pricing, and Decentralization

decisions come from top executives (centralization). For example, variable cost transfer prices allow centers to decide whether to buy from or sell to an internal or external party, based on improvements in their center’s cost or revenue; negotiated transfer prices allow firms to revise their production strategies to allow their own centers to create and capture more value, which will be reported in the center’s profit.

When should top management hand decision-making power to subordinates who are directing centers? When doing so allows the people with the best information to make decisions. Michael Jensen and William Meckling argue (link) that there are two types of knowledge relevant in making decisions, specific knowledge and general knowledge:

We define specific knowledge as knowledge that is costly to transfer among agents and general knowledge as knowledge that is inexpensive to transmit. Because it is costly to transfer, getting specific knowledge used in decision-making requires decentralizing many decision rights in both the economy and in firms.

Jensen and Meckling go on to compare the value of decentralization in the firm to the benefits of free market economies. They quoted the famous economist Frederick Hayek (link):

If we...agree that the economic problem of society is mainly one of rapid adaptation to changes in the particular circumstances of time and place....decisions must be left to the people who are familiar with these circumstances, who know directly of the relevant changes and of the resources immediately available to meet them. We cannot expect that this problem will be solved by first communicating all this knowledge to a central board which, after integrating all knowledge, issues its orders. We must solve it by some form of decentralization (Hayek 1945, p. 524).

These arguments imply that management should allow the directors of responsibility centers to make sourcing decisions and negotiate prices themselves when those directors have a great deal of specific knowledge that (by definition) they cannot easily transfer to their superiors. Rapid changes and wide variations in market conditions or production costs are hard to communicate, and should lead top management to shift price-setting rights to the directors of the centers. However, if appropriate transfer prices are based only on general knowledge, there is no reason for top management to delegate price setting to the directors. Delegation is also unwise if top management has specific information about the organization’s strategy, or about the overall cost of available resources (such as cash or investments in fixed assets), that is hard to communicate to their subordinates.
Approaches to Budgeting

Why Budget?

A **budget** is a financial plan that projects financial statement accounts, operational performance, and expenditures for one or more reporting periods. A sophisticated budget includes a complete set of financial statements that the firm would project if all of the operational and financial projections were realized. It also keeps track of the cash balance throughout the period (to make sure it is adequate).

The budget serves all four functions of financial reporting systems:

- **Attention Directing.** The budget process itself forces managers to spell out their strategic and financial plans, what they require to accomplish those plans, and what the impacts of those plans will be on financial position and operating performance. A finalized budget forces managers to explain deviations of actual from budgeted outcomes (**budget variances**).
- **Decision Facilitating.** The budget helps a firm to plan its resource expenditures and uses, to define expectations, and to focus its responses to deviations from the plan.
- **Decision Influencing.** The budget provides a basis for performance evaluation and compensation, by setting out a baseline against which outcomes can be compared.
- **Coordination Facilitating.** The budget helps the firm devise operational plans and resource allocations to balance the output of each organizational unit with the demands of its internal customers.

The power to serve all four of these functions makes the budget one of the single most important elements of a managerial reporting system in a decentralized firm.

Elements of a Budget

A comprehensive budget comprises four elements that each must be feasible, in light of the others:

- **Operational projections** describing the number of units to be produced, sold, and drawn from or added to inventory, as well as the resources required to achieve that proposed production and sales level, including personnel, raw materials, equipment, and other inputs.
- **Financial projections** describing the financial statements and key performance indicators that would result from the operational projections.
- **Cash-flow budgets** describing cash inflows, outflows, and balances at each point of the year, with special attention to whether the firm will have the cash on hand needed to meet demands...
on every day of the year. A firm that has insufficient cash resources on January 31 will fail, no matter how rosy the cash position will be in late autumn.

- **Performance standards** stating expected levels of efficiency, margin, and other aspects of performance for each worker or organizational unit. Performance standards generally state both a level of performance that must be achieved and a horizon at which performance will be measured. For example, the operational and financial projections may assume that an assembly line can produce 100 widgets every day, 3,000 widgets every 30 days, or 9,000 widgets every 90 days. These standards all reflect the same level (a rate of 100 widgets/day), but differ in horizon length.

### Methods of Budgeting

The vast majority of organizations develop a **periodic budget** based on a quarter, month, or year. Each periodic budget picks up where the prior one left off. For example, a firm might have one budget for the 2012 calendar year (ending December 31, 2012) and another that starts on January 1, 2013, for the 2013 calendar year. For many large firms, developing an annual budget takes five or six quarters (e.g., starting in July or September 2012 for the 2014 calendar year); budgeting is a long and painful process.

Some organizations attempt to streamline their budget process by using **rolling budgets** that overlap. For example, a firm might have a rolling 12-month budget for the 2012 calendar year, which would be updated at the completion of the first quarter, resulting in a 12-month budget from April 1, 2012, to March 31, 2013. The overlap in the budget periods makes it easier to update projections (since only 3 months have passed), and also gives the firm more rapid feedback on the quality of their budget process.

One shortcoming of rolling budgets is that they make it difficult for top management to usher in sweeping changes. Instead, management risks **budget creep**, in which every part of the organization demands slightly more money than they received in the previous budget cycle—whether they truly need it or not. While budget creep is a problem with periodic budgets as well, rolling budgets must build in a periodic element (with a more thorough review every twelve months, for example) to provide the time needed to reconsider the status of every unit and project.

One way to address budget creep is to implement **zero-base budgeting**, which requires everyone submitting a budget request to justify that request relative to a baseline of no funding at all, rather than the more common baseline of the previous budget allotment. Zero-base budgeting has been popular among many governmental and nonprofit organizations, because it suggests a fierce attitude of cost-cutting.
President Jimmy Carter argued the benefits of zero-base budgeting in a 1977 memorandum (link) directing all executive departments and agencies to use a zero-base budgeting system:

During the campaign, I pledged that immediately after the inauguration I would issue an order establishing zero-base budgeting throughout the Federal Government. An effective zero-base budgeting system will benefit the Federal Government in several ways. It will

- Focus the budget process on a comprehensive analysis of objectives and needs
- Combine planning and budgeting into a single process
- Cause managers to evaluate in detail the cost-effectiveness of their operations
- Expand management participation in planning and budgeting at all levels of the Federal Government

Zero-base budgeting often effectively forces managers to defend their expenditures with more than simply a nod to the status quo. However, its effectiveness in cutting spending is limited by the fact that, to the extent that it truly ignores prior budget allotments, it throws out very useful information; to the extent that it implicitly incorporates prior budget allotments, it differs little from traditional periodic or rolling budgets. (See this link for more on zero-base budgeting.)

Supply and Demand in the Budget Process

Balancing Supply and Demand for Output

One of the most important dimensions of coordination is balancing the firm’s production with market demand. Consider a very simple business with two organizational units: Production and Sales. In a well-coordinated business, Production produces approximately as much as Sales can sell. The budget process typically provides an iterative mechanism for Production and Sales to coordinate their efforts. First, Production shares information about how much it can produce in the coming period, and Sales shares information about how much it can sell. If Sales thinks it can sell more than Production can produce, the firm must decide whether to give Production more resources or to encourage Sales to trim its targets. If Production thinks it can produce more than Sales can sell, the firm must decide whether to lower Production’s targets or increase Sales’s targets.

A budget process that places priority on what Sales can sell (rather than what Production can produce) is often called pull budgeting, because the firm is effectively letting demand from customers “pull” production out of the company. Pull budgeting helps the firm avoid excess inventory, because it forces Production to produce no more than can be sold (plus any desired increase in inventory). Pull budgeting is widely viewed as superior to push budgeting, in which Production produces what it can and forces Sales to find a way to sell it, because forcing unwanted products and services on customers is thought to be about as effective as “pushing on a string.” However, the superiority of pull budgeting is hardly absolute: a firm with strong productive capacity must do everything it can to encourage stronger sales efforts.
Balancing Supply and Demand for Performance

Just as businesses must balance supply and demand for productive output, they must balance supply and demand for performance—especially financial performance. It is typically during the budget process that firms set performance targets. How hard should employees be working? How much should they be able to accomplish with the resources they are given? What returns should the firm provide to its investors? Just as a firm must balance the output Sales can sell with the output Production can create, it must coordinate the performance top executives and investors demand with the performance that employees can achieve.

Ideally, performance targets incorporate information drawn from lower-level managers (who are likely to know the most about their capacity, abilities, and market conditions), along with information on the needs of investors and upper management. But as a practical matter, budgets reflect demand and supply unequally. A budget that places priority on the demands of investors and upper management is typically called a top-down budget, while a budget that places priority on the abilities of employees is typically called a bottom-up budget.

Setting Performance Standards

Performance standards are the baseline against which performance is evaluated and compensated. Firms usually follow one of three philosophies in setting targets, which reflect the relative importance top management places on motivating employees, responding to budget variances, and planning in the face of uncertainty.

- Firms that place high value on motivation are likely to set stretch standards, defined as challenging levels of performance that are attainable, but more likely to be missed than exceeded. Academic research shows that people work hardest when presented with stretch standards.
- Firms that place high value on responding to budget variances are likely to set practical standards, defined as levels of performance that are expected, and as likely to be missed as exceeded. Practical standards allow managers to save time by devoting their attention to the relatively few line items with large budget variances (while largely ignoring aspects of performance in line with the budget), a practice known as management by exception.
- Firms that place high value on planning are likely to set minimum performance standards, defined as levels of performance that the firm can be quite certain of at least meeting, and probably of exceeding. Minimum performance standards help in planning, because unfavorable surprises disrupt planning more than favorable surprises. For example, assume a firm is relying on one division to bring in enough cash to support the financial commitments made by a second division. If the first division brings in too little cash, the firm might go bankrupt. If the first division brings in more cash than expected, however, the firm may not have acted optimally, but it will survive to try again another day. This is one reason for the popularity of the adage “underpromise and overdeliver.”
Approaches to Budgeting

Note that each type of performance standard sacrifices one goal in pursuit of another. Firms that employ stretch standards sacrifice planning and management-by-exception in exchange for motivating employees more effectively. Firms that employ practical standards sacrifice motivation and planning for the ease of management-by-exception. Firms that employ minimum performance standards sacrifice motivation and management-by-exception to be confident that their plans will be realized.
Paying Managers to Lie

Economist Michael Jensen published an article in the *Harvard Business Review* entitled “Corporate Budgeting Is Broken: Let’s Fix It ([link](#)).” Jensen argues that the primary goal of a budget should be to collect the best possible information from managers, to allow the firm to plan effectively. However, if the budget is also used as a baseline to evaluate performance, managers have a strong incentive to understate their ability to perform and overstate the resources they need to reach a targeted performance level, distortions commonly referred to as sandbagging. The goal of sandbagging is to generate budgetary slack in the form of (1) targets that are easy to outperform, (2) excess resources that allow better performance (empire building), or a combination of the two. Sandbagging arises because the *coordination-facilitating function of budgeting directly conflicts with its decision-influencing function*. As Jensen indicated with the original title of his article, using the budget to accomplish both of these goals is effectively “Paying Managers to Lie.” The solution, he argues, is to completely separate compensation policy from the budget.

*Kinks in Compensation Plans*

Jensen raises a second concern about linking compensation to budgets: budget-based compensation distorts managers’ operational decisions by making the pay-performance relationship even less linear that it would otherwise be. The figure below depicts a typical budget-based compensation plan.
The worker is guaranteed a minimum level of compensation (a floor) even if performance is low. If performance crosses a hurdle (usually based on a budget target), the manager gets a hurdle bonus, and an additional bonus that increases with performance up to a ceiling or cap. The compensation plan has several kinks where incentive intensity (the slope of the compensation function) changes dramatically. Incentive intensity is zero below the hurdle, but becomes infinitely high at the hurdle, where an infinitesimally small improvement in performance can earn the manager a hurdle bonus. Incentive intensity is moderate from the hurdle up until the ceiling is reached, at which point incentive intensity again drops to zero.

These changes in incentive intensity encourage managers to distort business decisions in ways that may not be in the firm’s interest. A salesperson who is already performing well enough to hit the compensation ceiling has an incentive to push additional sales into the next period. The additional sales won’t increase compensation in the current period, but might affect the next period’s bonus. The same is true for a salesperson who is performing so poorly that a hurdle bonus is unattainable: the salesperson has an incentive to reduce performance even more by pushing sales into the next period when they might increase the bonus. A salesperson in between those extremes has an incentive to pull sales from future periods into the current period to earn a higher bonus. In each of these cases, the kinks in the compensation plan give the salesperson an incentive to shift the timing of performance in ways that are unlikely to be in the firm’s best interest. Of course, the salesperson might choose not to exploit these opportunities, particularly if decreasing current performance might lead to termination. However, that does not change the fact that the kinks encourage the manager to consider actions that are not optimal for the firm.

Convexity and Concavity
More generally, kinks cause the compensation plan to be convex in some ranges and concave in others, as shown in the figure below. Assume that the manager can choose between achieving an average level of performance every period, or oscillating between high and low performance from one period to the next. When compensation is convex (meaning that incentive intensity increases with performance), the manager gains more by being above average in one period than she loses by being below average in the next. She will therefore avoid smoothing performance over time, preferring oscillation. When compensation is concave (meaning that incentive intensity decreases with performance), the manager loses more by being below average in one period than she gains by being above average in the next. She will therefore attempt to smooth performance over time.

Convex and concave compensation distorts risk-taking in similar ways. A manager with convex compensation seeks risk, because the benefit of good luck is greater than the cost of bad luck. A manager with concave compensation avoids risk because the cost of bad luck is greater than the benefit of good luck.

**Jensen’s Dream Contract**

Jensen concludes his article by advocating what I call Jensen’s Dream Contract, shown below. First, the contract does not incorporate any information from the budgeting process, so the firm is not paying managers to lie. Second, the contract does not have any kinks, curves, caps, or hurdles, so it doesn’t encourage managers to distort their risk-taking or performance-smoothing behavior.
Despite these two benefits, very few firms use Jensen’s Dream Contract outside of sales forces paid on straight commission. Why is that? One reason is that Jensen’s solution is extremely one-sided. In an attempt to use the budget as a more effective coordinating tool, his contract eliminates the decision-influencing value of the budget. In particular, the budget is very useful in tying compensation to items under managers’ control (the controllability principle), and in eliminating uncontrollable factors as much as possible. Even if managers’ forecasts cannot be trusted completely, any information on market conditions, production challenges, and other uncontrollable influences allows compensation to reflect managers’ contribution to performance through working harder, working smarter, and making intelligent decisions. Adhering more closely to the controllability principle reduces the risk that managers face, and therefore reduces the compensating differential managers must be paid to take a job with uncertain compensation.

Another problem with divorcing compensation from the budget process is that it isn’t clear how firms would set the terms of the contract without any budget input. The intercept and slope of the compensation plan must be set so that employees expect to be paid enough that they will take the job (the participation constraint). However, the firm doesn’t want to pay more than it needs to. How will the firm determine expected pay levels without using budgetary forecasts?

Finally, the straight-line contract may avoid the distortions that arise from convexity and concavity, but it does so by imposing a great deal of risk on the employee. The contract mimics the compensation paid to an entrepreneur—a straight proportion of the profits of the firm. But most employees are not entrepreneurs because they are risk-averse and value the security of a high minimum salary. Forcing a risk-averse worker to take an entrepreneur’s compensation plan requires large compensating differentials that may be more costly than smoothing, risk-taking or sandbagging.

The Solution

Jensen’s Dream Contract does not solve the problem he identifies, because changing the parameters of a contract doesn’t address the fundamental conflict of using the budget to both facilitate coordination and influence decisions. The problem is similar to one we saw in the essay “Paying for Performance,”
which pointed out the tension between wanting to motivate people by high incentive intensity and wanting to reduce compensating differentials by lowering workers’ exposure to risk. The solution to that tension was to invest in monitoring systems that would allow the organization to measure more precisely the aspects of performance under the workers’ control, allowing more incentive intensity without imposing more risk on the worker. (They could simply work harder to ensure higher pay.)

The solution to the problem Jensen identifies is similar: avoid paying managers to lie by investing more in systems to monitor and reduce both sandbagging and real measure management. Make managers aware that the accuracy of their budget forecasts will be tracked and inaccurate projections punished. Identify the most harmful forms of real measure management (such as accelerating or delaying sales), and let managers know that such actions will be tracked and punished. Such monitoring systems are costly, but often less costly than distorted budget forecasts and rampant measure management.

Don’t hold it against Jensen that he didn’t see this solution. After all, he is an economist, not an accountant!
The Theory of the Firm, Technology, and the Future of Work

IT giant IBM told Personnel Today that the firm's global workforce of 399,000 permanent employees could reduce to 100,000 by 2017, the date by which the firm is due to complete its HR transformation programme.

Tim Ringo, head of IBM Human Capital Management, the consultancy arm of the IT conglomerate, said the firm would rehire the workers as contractors for specific projects as and when necessary, a concept dubbed “crowd sourcing.”

“There would be no buildings costs, no pensions, and no healthcare costs, making huge savings,” he said.

Outsourcing experts said employers from both the private and public sector were increasingly using the model as they looked to squeeze people costs post-recession.

When asked how many permanent people IBM could potentially employ in 2017, Ringo said: "100,000 people. I think crowd sourcing is really important, where you would have a core set of employees but the vast majority are subcontracted out."

He stressed the firm was only considering the move, and was not about to cut 299,000 jobs, as staff would be rehired as contractors.

Source: Personnel Today, April 23, 2010 (link)

Introduction

Why would IBM reduce its headcount by 75%? Given that they would rehire many former employees as contractors for specific projects, would these changes actually make IBM “smaller,” or is the difference only one of legal form rather than economic substance? How would IBM’s reorganization affect those who remain employees? Those who are rehired as contractors? Finally, if this is a good idea now, why wasn’t it a good idea ten years ago? What has changed?

To answer these questions, this essay provides the tools needed to talk carefully about what it means to describe IBM as “an organization” (or, as most economists say, a “firm”), what it means for a worker to be an employee rather than an independent contractor, and what factors affect the wisdom of hiring each type of worker. As we will see, technology has strongly tilted the balance in favor of outside contractors, in a trend that will dramatically affect the careers of many business professionals.
Firms, Employees, and Independent Contractors

One of the most influential theories of the firm focuses on two features of organizations: contracts and transaction costs. Jensen and Meckling (link) defined the firm as a nexus of contracts. In this view, the firm is a legal entity that can be a party to contracts with a set of investors, suppliers, customers, and co-producers (as in the case of strategic alliances). Contracts with investors allow the organization to receive capital in exchange for promises to generate a return, which it does by using its capital to write contracts to secure the use of fixed assets, raw materials, and essential services.

The last century has seen an explosion in the forms of contracts organizations use. Firms can raise capital by issuing debt contracts, equity contracts, or contracts that mix the features of both (such as contingent convertible bonds (link)). Firms can buy fixed assets or lease them under a variety of terms, including synthetic leases (link) that, again, mix the features of both. Firms can secure essential services by hiring employees or outsourcing the service to people outside the organization. Contracting choices for capital and fixed assets are often motivated primarily by the details of tax and financial reporting regulations, and often have limited impact on how the organization actually operates and what types of managerial reporting systems best achieve its goals.

In contrast, contracting choices for services have dramatic effects on how organizations operate and design managerial reporting systems, because it is exactly these features that determine whether a person who provides services is an employee of the organization or an independent contractor to whom the service has been outsourced.

US tax law determines the status of a service provider based on the degree of control the organization exercises over the person providing the service, and the degree of independence the person retains over his or her actions. The US Internal Revenue Service (quoted from this link) looks to three sets of facts to determine control and independence:

- **Behavioral**: Does the company control or have the right to control what the worker does and how the worker does his or her job?
- **Financial**: Are the business aspects of the worker’s job controlled by the payer? (These include things like how worker is paid, whether expenses are reimbursed, who provides tools and supplies, etc.)
- **Type of Relationship**: Are there written contracts or employee-type benefits (i.e., pension plan, insurance, vacation pay, etc.)? Will the relationship continue, and is the work performed a key aspect of the business?

A typical employment contract gives the firm substantial behavioral control over the employee, makes the firm financially responsible for providing the resources needed to perform the job, and establishes a long-term relationship in which the employee depends on the employer for benefits and job security. The typical independent contractor agreement includes minimal behavioral control, focusing instead on the deliverable the contractor is to provide; it makes the contractor responsible for providing the resources necessary to produce that deliverable; and it establishes a short-term relationship that leaves
The Theory of the Firm, Technology, and the Future of Work

The contractor free to contract with others. (Because the employee-contractor distinction has many tax implications, many businesses attempt to craft contracts that give them the substantive benefits of one classification and the tax benefits of another. This essay focuses only on the substantive benefits.)

The Employment Contract

The most common employment contract is an at-will employment contract, which allows either party to cancel the contract on short notice for any reason. The freedom to dissolve the contract at short notice allows the organization to write contracts that impose a high degree of behavioral control, determining where and when employees work and specifying duties so broadly that it greatly limits the worker’s independence. These characteristics are illustrated in the at-will contract template provided by Inc.com. The bolded phrases in these slightly revised contract terms indicate the high degree of behavioral control and the broad duties required of the employee.

Places and Hours of Employment: Employee agrees that their duties shall be primarily rendered at Employer’s business premises during business hours or at such other places and times as the Employer shall in good faith require. Full time service for the Employee is expected which requires a minimum of ___ hours per week, exclusive of vacation, or any other form of leave as described within this Agreement.

Essential Functions and Duties: The essential job functions or duties of this position are as follows…. Employee shall also perform such other duties as are customarily performed by other persons in similar such positions, as well as such other duties as may be assigned from time to time by the Employer.

Duty of Loyalty and Best Efforts: Employee shall devote all of his/her working time, attention, knowledge, and skills to Employer’s business interests and shall do so in good faith, with best efforts, and to the reasonable satisfaction of the Employer. Employee understands that they shall only be entitled to the compensation, benefits, and profits as set forth in this Agreement. Employee agrees to refrain from any interest, of any kind whatsoever, in any business competitive to Employer’s business. The Employee further acknowledges they will not engage in any form of activity that is likely to damage the reputation of the Employer or produce a “conflict of interest” with the Employer unless agreed to in advance and in writing.

The at-will contract template also gives the firm great flexibility in measuring performance and linking those measures to retention, promotion, and raises:

Performance Terms: Employee performance will be determined by the benchmarks and other performance terms set by the Employer, and that may be changed by Employer at any time. The Employee understands that failure to reach benchmarks or performance terms may result in reassignment, demotion, or termination. Employee further understands that reaching these benchmarks or performance terms constitutes
The Theory of the Firm, Technology, and the Future of Work

a reasonable and substantial condition of employment but does not in any way guarantee or promise continued employment.

Like any arm’s-length transaction, the at-will contract between the employer and employee reflects a balance of interests. The contract provides the firm with a great deal of behavioral control over where, when, and how the employee works, what he or she does, and how that behavior is compensated. In exchange, the employee has a reasonable degree of long-term employment security, because although the firm can fire him or her at will, it is unlikely to do so unless the employee performs poorly. Finally, the firm bears the financial burden of giving the employee the resources required to perform the job, such as an office and desk, tools and equipment, and supporting services. The contract can be sweeping and rely heavily on discretion because either party can dissolve the contract on short notice.

The Independent Contractor Agreement

While employment contracts are sweeping and discretionary, independent contractor agreements are narrow and explicit. The independent contractor agrees only to provide a particular set of services, with compensation determined by outputs and outcomes. An agreement with an independent contractor provides the organization with no ability to control the behavior of the worker, but simply specifies the outputs or outcomes to be delivered to the firm in exchange for compensation (deliverables). As the IRS summarizes it (link),

The general rule is that an individual is an independent contractor if the payer has the right to control or direct only the result of the work and not what will be done and how it will be done.

The Chamber of Commerce (link) provides a template for an independent contractor agreement. Here are some of its key provisions, verbatim.

Duties, Term, and Compensation. The Contractor’s duties, term of engagement, compensation, and provisions for payment thereof shall be as set forth in the estimate previously provided to the Company by the Contractor and which is attached as Exhibit A, which may be amended in writing from time to time, or supplemented with subsequent estimates for services to be rendered by the Contractor and agreed to by the Company, and which collectively are hereby incorporated by reference.

Expenses. During the term of this Agreement, the Contractor shall bill and the Company shall reimburse [him or her] for all reasonable and approved out-of-pocket expenses which are incurred in connection with the performance of the duties hereunder. Notwithstanding the foregoing, expenses for the time spend by Consultant in traveling to and from Company facilities shall not be reimbursable.

Written Reports. The Company may request that project plans, progress reports, and a final results report be provided by Consultant on a monthly basis. A final results report shall be due at the conclusion of the project and shall be submitted to the Company in a confidential written
The “Exhibit A” referred to in the contract template would need to spell out the required product or service in considerable detail, because neither party can dissolve the contract unilaterally—they would need to renegotiate and come to a new, mutually acceptable agreement.

Transaction Costs and the Size of the Firm

Armed with our new definitions, we can begin to analyze the effects and wisdom of IBM’s plan to replace employees with independent contractors. IBM would sacrifice the ability to control how the people they contract with provide essential services, the flexibility of being able to redefine job requirements and compensation, and the ability to control behavior outside the very narrow scope of the deliverable being provided. In exchange, they shift the financial burden of providing the resources required to perform the service onto their independent contractors.

When are the benefits of independent contractors greater than the costs? We can draw some conclusions from considering the difficulty of writing contracts to specify the deliverables an independent party would be asked to provide, and the nature of the investment required to provide those deliverables.

Contracting Perspective

Ronald Coase (1937) analyzed the employee-contractor decision in terms of the costs of writing and monitoring contracts, which he termed transaction costs. One benefit of a flexible and broad at-will contract is that the firm doesn’t need to write many of them. Most employees at a firm sign the same at-will employment contract, and do so only once, because the firm can make unilateral revisions to job duties and performance evaluation as circumstances demand. In contrast, a firm would need to write a different contract for each type of service it wishes to secure through an independent contractor, because each contract needs to specify a different set of requirements and payments. Moreover, because the contract is so explicit, changes in market conditions or the firm’s needs entail writing a new contract. From a transaction costs perspective, independent contractors are more desirable than employees when the deliverables required of service providers are more easily measured, standardized, and granularized:

- **Measurability** of deliverables. Measurability captures the ease with which a firm can describe an objective and legally enforceable set of proxy measures upon which payment for a service can be based.
- **Standardizability** of deliverables. Standardizability captures the firm’s ability to reuse the terms specifying the deliverables required of a contractor.
- **Granularity** of deliverables. Granularity captures the firm’s ability to carve out a limited set of deliverables that can be performed independently of any others.
For example, IBM might find it relatively easy to terminate employees responsible for providing support to clients using IBM software, and then retain those services through independent contractors. IBM could measure delivered performance almost entirely through a combination of outputs (hours spent with customers) and outcomes (customer satisfaction scores), and do so through a standard contract applied to all independent contractors providing support across a wide range of business segments. The customer-support contractors can also perform effectively without interacting with any of IBM’s other employees or contractors.

IBM would probably have a harder time outsourcing the task of pricing consulting contracts. Success in pricing can’t be measured effectively until after the work is finished, at which time IBM still has only rough, imperfect information about the true costs of the engagement. Contracts for pricing would be difficult to standardize, because the time and difficulty of the task depend on market conditions and the specifics of the engagement, which vary significantly from month to month and engagement to engagement. Finally, pricing decisions are not easily granularized, because pricing decisions are tightly tied to other strategic decisions, such as investment in resources, overall product mix, and market focus.

Investment Perspective

The best answer to the employee-contractor decision is also determined how much investment would be required of an independent contractor, and how specifically tailored to the employer those investments must be. In particular, the firm must consider:

- **Capital intensity** of investment. Capital intensity measures the dollar value an independent contractor would need to invest per dollar value of labor provided.
- **Specificity of investment.** Specificity measures the ability of an independent contractor to repurpose their capital investment to be an effective contractor for a different organization.

Traditional employment is a very low capital-intensity proposition, because the employer typically bears all of the costs of providing buildings, equipment, and hardware. (Common exceptions are personalized tools, as in the case of carpenters, and intellectual capital, such as scientific or professional training.) In contrast, contractors are required to provide all of their own resources, including offices and production facilities, computers, and support staff. High capital intensity makes independent contracting less attractive, especially if the firm can provide those investments at lower cost due to economies of scale.

The impact of capital intensity depends on how specifically the investments are tied to an employer. Assume that a contractor who repairs equipment needs to make a large investment in machine-specific maintenance tools and training. Investment of this sort makes independent contracting especially undesirable if only one firm uses the machines (a highly specific investment). The contractor suffers much the same lack of independence as an employee, because she has few options to contract with other businesses. If many firms use the machines, however, the contractor has many options to earn a return on her investment; she could, for example, work with a business and its competitors, a form of independence explicitly prohibited by most employment contracts.
Technology and the Future of Work

If it makes sense today for IBM to shed 75% of its employees and rehire many back as independent contractors, why didn’t it make sense 10 years ago? Yogi Berra tells us that predictions are hard, especially about the future, futurist Margaret Regan sees the future of work this way (link):

But nothing will compare to the surprises waiting for you in 35 years. By 2042, there will be no 'workplace' as we know it. Goodbye wired desks; hello wireless personal wearable networks. It will be difficult to even talk about your "dream job," since 70 percent of the job types you might pursue do not exist today, and 80 percent of the technology is not invented yet.

Your younger colleagues will have 10 careers in their lifetime, so loyalty will be to the profession, not to the employer. Work will feel like a Hollywood movie in which managers, directors, and team members come together around a specific project, collaborate intensely, and then disband. The new employment contract will say, 'I will stay as long as it's a good deal for me.'

Although Regan might be overstating her case, technology has dramatically shifted contracting and investment costs in favor of employment-free independent contracting.

Investment Perspective

Technology has had its most obvious impact on reducing the intensity and specificity of capital investment required for workers to provide useful services to large organizations. Before electronic communication, organizations were forced to communicate with their workers through paper and face-to-face meetings, so it was important to have a shared facility in which to store and distribute paper, do the work, and hold the meetings. Such facilities require a significant investment in shared resources, so traditional employment (with heavy employer investment) was a natural solution. Ubiquitous wireless networks and smartphones now allow workers to communicate without either the firm or the worker needing to invest in office buildings, relying instead on very cheap devices that can access Web sites and support videoconferencing. Minimal investment now allows an independent contractor to communicate very effectively with a large organization.

Technology has also reduced the specificity of the investments required of independent contractors. When technology was new, organizations often purchased or developed their own idiosyncratic software systems, forcing a contractor to invest in time, software, and hardware specific to each company. With the maturing of the software industry, a small number of software packages and communication protocols can serve most large companies, allowing contractors to invest their resources in expertise and hardware that allow them to serve entire industries.

Contracting Perspective

Technology has also made it easier to arrange contracts for measurable, standardized, and granular deliverables. The greater ease of contracting is perhaps best exemplified by the rise of crowdsourcing Web sites that allow large firms to buy many very small, highly granularized services from independent
contractors at very low per-unit costs. Amazon Mechanical Turk (link) is one of the best-known markets. As Amazon describe their “Market for Human Intelligence Tasks” (link),

Amazon Mechanical Turk is a marketplace for work that requires human intelligence. The Mechanical Turk service gives businesses access to a diverse, on-demand, scalable workforce and gives Workers a selection of thousands of tasks to complete whenever it’s convenient.

Amazon Mechanical Turk is based on the idea that there are still many things that human beings can do much more effectively than computers, such as identifying objects in a photo or video, performing data de-duplication, transcribing audio recordings, or researching data details. Traditionally, tasks like this have been accomplished by hiring a large temporary workforce (which is time consuming, expensive, and difficult to scale) or have gone undone....

A Human Intelligence Task, or HIT, is a question that needs an answer. A HIT represents a single, self-contained task that a Worker can work on, submit an answer, and collect a reward for completing.

Other sites offering similar marketplaces are 99designs (link), which specializes in logo and graphic design contracts, and Coffee and Power (link), which connects individuals to individuals for small jobs.

All of these marketplaces rely very heavily on granularity. Harvard Law professor Yochai Benkler illustrates the importance of granularity by comparing the successful crowdsourced encyclopedia Web site Wikipedia to the far less successful crowdsourced textbook Web site, Wikibooks. Writing or editing an encyclopedia Web site is a highly granularized task—one need only have clear knowledge of the topic discussed in that entry, and the overall format of the Web site. There is no need to delve into any entries other than the one being edited. In contrast, writing or editing a portion of a textbook is far less granularized. An author would be ill-advised to work on Chapter 2 of a cost accounting book without knowing how Chapter 1 introduced the material and how Chapter 3 will carry it forward. As a result, many more people can contribute to Wikipedia with minimal investment, and be assured of a more successful product than WikiBooks can provide. (Here is a link to a review of Benkler’s book, The Wealth of Networks, which includes a link to a free version of the entire book.)

I’m an MBA. Is My Job at Risk?

Most MBA and EMBA students have executive positions. Is your job likely to be outsourced? The answer depends on how closely your job meets the precise definition of an executive function: a task that involves allocating costly resources (such as attention or investment) among separate granular tasks. To the extent that you are allocating resources among many different tasks, your job is not granular, and is not easily outsourced.

Gamification

Any discussion about the future of work would be incomplete without acknowledging the increasing popularity of gamification, by which firms incorporate popular aspects of games into the workplace.
Even without technology, it isn’t much of a stretch to see work as a game, in which the players are given challenging tasks to perform under significant time pressure, and receive rewards based on measured performance. In *Total Engagement* ([link](#)), authors Byron Reeves and Leighton Reed advocate modeling the workplace more closely on modern computer games, especially those that take place in virtual worlds (like World of Warcraft and Second Life). They argue that gamification makes work far more engaging because computer games provide rich and immediate feedback, a social environment that includes cooperation with allies as well as competition with others, rich systems for communication, extensive markets and economies in which players can trade resources and favors, and clear systems for building credentials (“leveling up”) and reputations. For more information on gamification, read Bloomberg Businessweek’s *CEO Guide To Gamification* ([link](#)), and watch this fascinating video lecture by Carnegie Mellon professor (and former Disney Imagineer) Jesse Schell ([link](#)).
Part VIII: Sniffing Out Efficiency
Introduction

Standard costing systems report the costs of products and services according to the resources they should have consumed, given budgeted performance standards. This approach differs substantially from normal costing systems, which report costs according to the resources actually consumed. (Don’t forget that accountants use standard and normal in very technical senses. Both terms are synonyms for typical or common in ordinary conversation, and indeed both types of costing systems are fairly typical and common.)

Standard costing systems were developed and refined during the industrial era, and they are among the most important developments in recent accounting history. These systems integrate the two most useful and time-consuming elements of management reporting systems, the budgeting process and the cost accounting system, to highlight how actual performance differs from the assumptions underlying the budget. Furthermore, they all firms to construct detailed financial statements in which a single employee “owns” every line item and is held responsible for knowing its status, explaining why performance deviated from the budget, and proposing responses. Standard costing systems are useful mainly for “process shops” engaged in mass production, rather than “job shops” that provide customized services. This is because it is easier to devise standards when productions processes are repeated over and over with minimal changes, and because saving money on such a process benefits the firm every time it is repeated. (See the essay “Different Costs for Different Purposes” for more discussion of this topic.)

To see the differences between standard costing and normal costing, consider Careful Cookies, a fictional company consisting of only two departments, Mixing and Baking. The Mixing department’s performance standards dictate that it should cost $30 in labor to mix the dough for a batch of cookies, but for one batch the department actually uses $40 of direct labor. A normal costing system would apply $40 of labor to the batch, while a standard cost system would apply only $30 of labor (what it should have cost) to Work-In-Process inventory. The remaining $10 of cost is recorded as a variance, which we include as a cost-increasing adjustment to Cost of Goods Sold. Similarly, if the batch of cookies required only $25 of labor, $30 would still be applied to Baked Cookies inventory account, along with a $5 variance that is reported as a cost-reducing adjustment to Cost of Goods Sold.

Standard costing techniques meet four goals at once, all natural consequences of the structure imposed by double-entry bookkeeping:
1. **Standard costing purges an account owned by one party of variances caused by another.** Under standard costing, the Mixing department always transfers $30 of cost to the Baking department, regardless of the actual mixing cost. Deviations from the performance standards for Mixing are not Baking’s responsibility, and are not reported on Baking’s financial statements.

2. **Each party remains responsible for their own variances.** Since cost overruns or underruns are not passed on to the next department, where are they reported? In the financial statements of the party responsible for the overrun (Mixing), as part of a variance account!

3. **Variances are reported on the income statement during the period in which they arise.** Under normal costing, the costs or benefits of deviating from a performance standard flow through a series of inventory accounts on the balance sheet, and do not reach the income statement until products or services are sold at a higher or lower cost than the firm’s performance standards specify. When a cost overrun occurs, a normal-costing firm reports fewer raw materials (or more liabilities) than performance standards dictate, but reports more work-in-process or finished goods. As a result, the wealth reported on the financial statements is completely unaffected by the deviation. In contrast, a standard-costing firm records a variance as an expense or loss on the income statement as soon as production occurs, causing an immediate and (and very noticeable) reduction of wealth.

4. **Cost overruns are not reported as assets.** The conceptual definition of an asset requires that an item have a probable future economic benefit for the firm. By recording cost overruns on the income statement as expenses or losses, the firm avoids treating wasted money as if it represents a future benefit. By the same token, by recording unexpected efficiencies as gains (or reductions in expenses), the organization avoids reporting lower production costs as if they represent a reduction in future gains (lower assets).

Subsequent essays discuss how to calculate standard costs and variances, and how to record them using double-entry bookkeeping. The remainder of this essay focuses on the reports generated by standard costing systems, and explains how they are useful.

### Cost of Goods Sold, Revisited

Recall that normal costing systems distinguish between two types of Cost of Goods Sold (CGS). Unadjusted CGS is simply the sum of the reported cost applied to every item sold. Adjusted CGS adjusts that amount to reflect differences between the total costs applied and the total costs incurred during the period. Here is a sample Cost of Goods Sold statement for our fictional cookie producer, Careful Cookies:

<table>
<thead>
<tr>
<th>Cost of Goods Sold Statement (Normal Costing)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted Cost of Goods Sold</td>
<td>100,000</td>
</tr>
<tr>
<td>Adjustment for Over- or Underapplied Overhead</td>
<td>12,000</td>
</tr>
<tr>
<td>Adjusted Cost of Goods Sold</td>
<td><strong>112,000</strong></td>
</tr>
</tbody>
</table>
Isolating Responsibility Through Standard Cost Reporting

This statement isn’t very useful for assessing production efficiency, because the adjustment isolates only one deviation from expectations—in this case, more overhead was incurred than applied to the units sold. But over- or under-application of overhead is not necessarily a sign of the firm’s efficiency. They might have simply produced more or less than projected by the budget to meet demand. Moreover, Unadjusted Cost of Goods Sold doesn’t provide any information about how the incurred direct costs might have deviated from what they should have been—all of those direct costs are included in Unadjusted Cost of Goods Sold, even if they represent cost overruns.

Standard Costing generates a more informative report by separately reporting the amount that the sold goods should have cost as “Cost of Goods Sold @ Standard.” It further breaks the difference between standard cost and actual total cost into an adjustment for over- or underapplied overhead, and an adjustment for deviations of actual direct costs from standard levels:

<table>
<thead>
<tr>
<th>Cost of Goods Sold Statement (Standard Costing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted Cost of Goods Sold @ Standard</td>
</tr>
<tr>
<td>Adjustment for Over- or Under-applied Overhead</td>
</tr>
<tr>
<td>Direct Cost Variances</td>
</tr>
<tr>
<td><strong>Adjusted Cost of Goods Sold</strong></td>
</tr>
</tbody>
</table>

Drilling Down into Direct Cost Variances

The Standard Costing report above still falls short of isolating responsibility, because the two adjustments to Cost of Goods Sold may have many causes, each the responsibility of a different person. Ideally, the adjustments would be broken down in such detail that each line item on the report can be “owned” by a single manager, who is responsible for knowing the current state of the item (are we doing better or worse than the standards laid out in the budget?), explaining the current state, and proposing a solution. A good standard costing system allows a manager to drill down into a variance to get more information. A drill-down analysis might result in the following report:

<table>
<thead>
<tr>
<th>Cost of Goods Sold Statement (Standard Costing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted Cost of Goods Sold @ Standard</td>
</tr>
<tr>
<td>Adjustment for Over- or Under-applied Overhead</td>
</tr>
<tr>
<td><strong>Cost of Goods Adjusted for Over- or Under-applied OH</strong></td>
</tr>
<tr>
<td>Direct Material Variances—Baking</td>
</tr>
<tr>
<td>Direct Material Variances—Mixing</td>
</tr>
<tr>
<td>Direct Labor Variances—Baking</td>
</tr>
<tr>
<td>Direct Labor Variances—Mixing</td>
</tr>
<tr>
<td><strong>Adjusted Cost of Goods Sold</strong></td>
</tr>
</tbody>
</table>
Isolating Responsibility Through Standard Cost Reporting

By isolating variances by resource (materials and labor) and by department (Baking and Mixing), this report potentially allows four managers to each take ownership of a different direct material variance (a manager in charge of materials and a manager in charge of labor in each department). Even this report would benefit from further drill-down analysis. For example, there might be three different shifts in the Baking department, each working on one of two different products. Each shift’s contributions to the 7,000 favorable (cost-reducing) variance differences in their labor variances can be identified by segregating the variances by a report such as this:

<table>
<thead>
<tr>
<th>Direct Labor Variances</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Labor Variance—Baking ChocChip Shift 1</td>
<td>(3,000)</td>
</tr>
<tr>
<td>Direct Labor Variance—Baking ChocChip Shift 2</td>
<td>2,000</td>
</tr>
<tr>
<td>Direct Labor Variance—Baking ChocChip Shift 3</td>
<td>(4,000)</td>
</tr>
<tr>
<td>Direct Labor Variance—Baking Oatmeal Shift 1</td>
<td>(1,000)</td>
</tr>
<tr>
<td>Direct Labor Variance—Baking Oatmeal Shift 2</td>
<td>1,000</td>
</tr>
<tr>
<td>Direct Labor Variance—Baking Oatmeal Shift 3</td>
<td>(2,000)</td>
</tr>
<tr>
<td>Direct Labor Variances—Baking</td>
<td>(7,000)</td>
</tr>
</tbody>
</table>

In this report, we see that Shift 3 accounts for fully $6,000 of the $7,000 favorable variance, and most of that ($4,000) arises when they are baking ChocoChip cookies. We can drill down one more step, by decomposing that $4,000 into two different ways spending on a variable-cost resource may vary from budgeted performance standards: the cost per hour of labor may differ from the standard cost per hour, and the number of hours used to make each batch of cookies may differ from the standard hours allowed for that level of production. This variance decomposition into price and quantity effects might result in a report like this:

<table>
<thead>
<tr>
<th>Direct Labor Variance Decomposition—Baking ChocChip Shift 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Labor Rate Variance</td>
<td>500</td>
</tr>
<tr>
<td>Direct Labor Efficiency Variance</td>
<td>(4,500)</td>
</tr>
<tr>
<td>Direct Labor Variances—Baking ChocChip Shift 3</td>
<td>(4,000)</td>
</tr>
</tbody>
</table>

The unfavorable (cost-increasing) labor rate variance shows that the department has paid more than the standard rate per hour for the workers, but perhaps it was worth it. Either through the motivating effect of higher pay or the choice to use higher-ability (but more expensive) workers, the favorable (cost-reducing) labor efficiency variance shows that that shift was able to get their work done with far fewer hours than allotted by the performance standard.

We can apply a similar analysis to direct materials. In the report below, the favorable direct materials price variance suggests that each unit of raw materials was purchased at a price lower than that set by the standard, reducing Cost of Goods sold by 500. However, the unfavorable direct materials usage variance shows that each batch of cookies used slightly more raw materials than the standard allowed.
Perhaps the department bought lower-quality ingredients, which resulted in more cookies failing to meet quality standards (increasing the direct materials used per batch).

**Drilling Down into Indirect Cost (Overhead) Variances**

The same drill-down principles used for direct costs can also allow for reports that clarify the causes of under- and overapplied overhead. The first stage of the drill-down analysis is to segregate variances by overhead pool. Assume, for example, that each department has two overhead pools: one for machine-related overhead and another for labor-related overhead in the Mixing department (with machine hours and direct labor dollars serving as drivers); and one for oven-related overhead and another for labor-related overhead in the Baking department (with oven hours and direct labor dollars serving as drivers). Then the first step in a drill-down analysis of overhead might look like this:

<table>
<thead>
<tr>
<th>Sources of Over- and Under-Applied Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine-Related Overhead Pool—Mixing Department</td>
</tr>
<tr>
<td>Labor-Related Overhead Pool—Mixing Department</td>
</tr>
<tr>
<td>Oven-Related Overhead Pool—Baking Department</td>
</tr>
<tr>
<td>Labor-Related Overhead Pool—Baking Department</td>
</tr>
<tr>
<td>Over- or Underapplied Overhead</td>
</tr>
</tbody>
</table>

Decomposing over- or underapplied overhead in or of any one of these overhead pools is more difficult than decomposing a direct cost variance, because a single overhead pool includes both variable and costs. We can decompose the variable costs into a price effect and a quantity effect just as we decomposed direct labor above. The *Variable Overhead Spending Variance* captures the difference between the predetermined overhead rate and the actual spending on variable overhead per unit of the overhead driver used (a price effect). The *Variable Overhead Efficiency Variance* captures the difference between the amount of the driver actually used, and how much of the driver should have been used according to budgeted standards.

The decomposition of fixed overhead is slightly different, because (by definition) producing more units does not cause fixed costs to increase. As a result, it doesn’t make sense to decompose fixed overhead into a price effect and a quantity effect. Instead, standard costing systems decompose over- or underapplied fixed overhead into a *Fixed Overhead Budget Variance*, which captures how the actual total fixed costs deviated from the budgeted total, and a *Fixed Overhead Volume Variance*, which captures how the actual production volume differs from the activity base assumed in the calculation of the Predetermined Fixed Overhead Rate. For example, assume that a department budgeted for $20,000
of fixed overhead, and chose an activity base of 1000 batches, for a predetermined fixed overhead rate of $20/batch. However, actual fixed costs were $22,000 and the department made only 750 batches. In this case, the FOH Budget Variance would be an unfavorable $2,000 (the difference between the $22,000 actual FOH and the budgeted $20,000). The FOH Volume Variance would be an unfavorable $5,000 (the difference between the $20,000 that would have been applied according to the activity base, and the $15,000 = $20/batch x 750 batches that was actually applied to inventory).

A comprehensive Cost of Goods Sold report that reports the eight different types of variances discussed so far might something look like this (taking liberties with the exact numbers):

<table>
<thead>
<tr>
<th>Cost of Goods Sold Statement (Standard Costing)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted Cost of Goods Sold @ Standard</td>
<td>80,000</td>
</tr>
<tr>
<td>Variable Overhead Spending Variance</td>
<td>5,000</td>
</tr>
<tr>
<td>Variable Overhead Efficiency Variance</td>
<td>(8,000)</td>
</tr>
<tr>
<td>Fixed Overhead Budget Variance</td>
<td>4,000</td>
</tr>
<tr>
<td>Fixed Overhead Spending Variance</td>
<td></td>
</tr>
<tr>
<td>Total Adjustment for Under- or Overapplied Overhead</td>
<td>11,000</td>
</tr>
<tr>
<td><strong>Cost of Goods Sold Adjusted for Over- or Underapplied Overhead</strong></td>
<td>92,000</td>
</tr>
<tr>
<td>Direct Material Purchase Price Variances</td>
<td>12,000</td>
</tr>
<tr>
<td>Direct Material Usage Variances</td>
<td>(8,000)</td>
</tr>
<tr>
<td>Direct Labor Rate Variances</td>
<td>10,000</td>
</tr>
<tr>
<td>Direct Labor Efficiency Variances</td>
<td>16,000</td>
</tr>
<tr>
<td><strong>Adjusted Cost of Goods Sold</strong></td>
<td>112,000</td>
</tr>
</tbody>
</table>

**Sale Variances**

Many firms use performance standards to provide detail about revenues as well as costs. For example, assume that Careful Cookies’s budget projected that they would sell 3,000 batches of cookies at an average price of $100 per batch, for a total of $300,000 of revenue. Their actual revenue could deviate from the budget because the actual selling price differed from $100, or because sales volume deviated from 3,000 units. The Revenue portion of a Standard Cost Income Statement might look like this:

<table>
<thead>
<tr>
<th>Revenue</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenues @ Standard</td>
<td>300,000</td>
</tr>
<tr>
<td>Sales Price Variance</td>
<td>40,000</td>
</tr>
<tr>
<td>Sales Volume Variance</td>
<td>(60,000)</td>
</tr>
<tr>
<td><strong>Adjusted Revenue</strong></td>
<td>280,000</td>
</tr>
</tbody>
</table>

The favorable (revenue-increasing) **Sales Price Variance** reflects the additional revenue brought in because sales prices were above the standard of $100/batch, perhaps due to an unwillingness to provide discounts. The unfavorable (revenue-decreasing) **Sales Volume Variance** reflects the revenue
shortfall arising from selling fewer batches of cookies. Overall, the benefit of higher prices was more than offset by the lower sales volume, resulting in total revenue below that projected in the budget.
The Central Mechanism of Standard Costing

The goal of standard costing is to provide a report in which each deviation from the budget is reported as a separate line item that a single manager can “own.” Such reports were described in the previous essay, “Isolating Responsibility Through Standard Cost Reporting.” This essay discusses the mechanics of the double-entry bookkeeping system required to generate these useful reports. The central mechanism of standard cost bookkeeping is this:

Standard cost systems reduce resource accounts by the recorded value of the resources actually consumed, but apply to inventory the cost of resources that should have been consumed according to the budget’s performance standards. Any difference between the two amounts is recorded in a variance account, which is ultimately reported as an adjustment to Cost of Goods Sold.

For example, assume a purchasing manager paid $1,000 cash to buy 200 pounds of chocolate, but the performance standard dictates that the cost of chocolate should be only $4/pound. The standard costing system requires taking a credit of $1,000 from the cash account, because otherwise our records of our cash balance would not agree with the bank’s records. However, the firm debits the Raw Materials account for only $800 (200 pounds x $4/pound). The $200 difference is reported as a debit to the Direct Materials Purchase Price Variance, as follows:

Dr. Raw Materials—Chocolate $800  
Dr. Direct Materials Purchase Price Variance $200  
Cr. Cash $1,000

To record the standard cost of materials purchased and the associated variance

Now, imagine that a production manager uses 3 pounds of chocolate for a batch of cookies that, according to the standard, should have required 3.5 pounds of chocolate. The bookkeeper must reduce the Raw Materials—Chocolate account by $12, because that is the cost at which the 3 pounds of chocolate are already recorded in the account (3 pounds @ $4/pound). However, the bookkeeper also applies $14 to the Work-In-Process Inventory account, because the batch should have used 3.5 pounds according to the performance standard. She reports the $2 difference as a credit to the Direct Materials Usage Variance account, as follows:

Dr. Work-In-Process Inventory $14  
Cr. Direct Materials Usage Variance $2

To record the material usage variance and the associated adjustment to Work-In-Process Inventory.
Dr. WIP—ChocoChip Cookies  $14
  Cr. Direct Materials Usage Variance  $2
  Cr. Raw Materials—Chocolate  $12

To record the standard cost of materials used in production and the associated variance

The two variances accounts are income statement accounts included as part of the Adjustment to Cost of Goods Sold. The Direct Materials Purchase Price Variance is reported as one line item in the adjustment, and it is the sole responsibility of the purchasing manager, because it reflects his decision to pay more for Chocolate than the budget specified, but it isn’t affected by how efficiently the Chocolate was used. The bookkeeper reposts the Direct Materials Usage Variance as a separate line, and it is the sole responsibility of the production manager, because it reflects the efficient use of Chocolate in the production department, and it isn’t affected by the price of the Chocolate.

Decomposing Material Costs into Price and Quantity Effects

The entries above isolate the responsibilities of the production and purchasing manager by recognizing that any variable cost (such as direct materials) is the product of the price and quantity of the resource consumed.

The entries capturing the separate price and quantity effects for Direct Materials can be written more generally by using the following notation:

- **AP** = Actual Price/unit (as assessed by purchase invoices)
- **SP** = Standard Price/unit (as specified in the budget)
- **AQ** = Actual Quantity (purchased or used in production)
- **SQ** = Standard Quantity Allowed for Production (units actually produced x standard quantity/unit as specified in the budget)

The two entries then become:

Dr. Raw Materials—Chocolate  SP x AQ Purchased
  Dr./Cr. Direct Materials Purchase Price Variance  (AP – SP) x AQ Purchased
  Cr. Cash  AP x AQ Purchased

To record the standard cost of materials purchased and the associated variance

Dr. WIP—ChocoChip Cookies  SP x SQ
  Dr./Cr. Direct Materials Usage Variance  (AQ Used – SQ) x SP
  Cr. Raw Materials—Chocolate  SP x AQ Used

To record the standard cost of materials used in production and the associated variance
Decomposing Variances

Direct materials variances are a special case, because two separate events determine the two different variances we record: we recognize the DM Purchase Price Variance when we purchase materials, and we recognize the DM Usage Variance when we use them. For most other resources, we are forced to decompose a single difference between costs incurred and costs applied into two separate variances. You will find it easier to follow and remember the analysis if you use the following simple graphical image, which I call the Variance Crown (to accompany the Queen’s Cross):

The Variance Crown

At the top left of the crown we write the cost incurred when we consumed a resource. At the top right we write the cost we applied. In the middle we write the dollar amount we will use to split the total difference into two variances. The accounts affected, the calculation of the split point, and the names of the variances differ depending on the resource in question, but the overall structure is the same for all of the resources we examine in the remainder of this essay.

Decomposing Direct Labor Variances

It is a little more difficult to decompose direct labor variances into price and quantity effects, because only one event (the use of direct labor) determines the price and quantity at the same time. To help yourself grasp the intuition, recall from your elementary-school math that the product of two numbers describes a rectangle, so the actual and standard cost of labor can each be represented as follows:
Rectangular Representation of Variable Costs

Actual Cost = Actual Price x Actual Quantity
= AP x AQ

Standard Cost = Standard Price x Standard Quantity
= SP x SQ

The total deviation between actual and standard cost is reflected in the difference in the area of the two rectangles. Laying one over the other identifies how much difference in their areas is driven by actual price deviating from standard price, and how much is driven by actual quantity of labor deviating from the standard quantity:

Rectangular Variance Decomposition

The area of the rectangle labeled “Pure Price Variance” reflects the extra cost generated purely by the fact that actual labor price/hour was higher than the performance standard specified. The area of the rectangle labeled “Pure Quantity Variance” reflects the excess costs generated by the department’s using more hours of labor than specified in the standard. The rectangle labeled “mix” reflects a combination of the two factors: its area is determined by the excess price and the excess quantity.
The rectangular analysis conveys a nice intuitive understanding of decomposing variable costs when both price and quantity are higher than standard, but it gets confusing otherwise. One problem is that when price or quantity is lower than standard, the areas of the variance rectangles are negative. That is not a situation that arises often in elementary-school math. Another problem is that the rectangular analysis doesn’t clarify who should own the “mix” of the variances—should it be assigned to the person responsible for prices or for quantities?

Instead of using rectangles, most firms use a clever algebraic trick to create two variances:

The total variance is

\[ \text{Variance} = (\text{AP} \times \text{AQ}) - (\text{SP} \times \text{SQ}) \]

Now, add and subtract the quantity \(\text{SP} \times \text{AQ}\), leaving the total difference unchanged:

\[ \text{Variance} = (\text{AP} \times \text{AQ}) - (\text{SP} \times \text{AQ}) - (\text{SP} \times \text{AQ}) - (\text{SP} \times \text{SQ}) \]

The product \(\text{SP} \times \text{AQ}\) serves as a “split point” against which the actual and standard cost can each be compared. Now we can simply rearrange terms and simplify:

\[ \text{Variance} = (\text{AP} \times \text{AQ} - \text{SP} \times \text{AQ}) - (\text{SP} \times \text{AQ} - \text{SP} \times \text{SQ}) \]

\[ = \text{AQ} (\text{AP} - \text{SP}) + \text{SP} (\text{AQ} - \text{SQ}) \]

\[ = \text{Price Variance} + \text{Quantity Variance} \]

The last equation divides the total variances into two parts by comparing both actual and standard cost to \(\text{SP} \times \text{AQ}\), a split point that reflects the cost that would have been incurred had the firm used the actual hours they did, but paid the standard cost per unit. Note that this choice leaves the manager responsible for setting prices with some responsibility for the quantity used, because a higher actual quantity increases the price variance. However, this error in isolating responsibility is usually quite small.

As an example of the decomposition, assume that the performance standard allows a labor price of $20/hour, and that each batch of cookies should take 5 hours of labor. Perhaps because a manager decided to hire less-qualified workers at a lower wage, the actual price of labor was $15/hour, but it took 12 hours to cook two batches of cookies.

The variances are calculated as:
Labor Price Variance = AQ \times (AP - SP) = 12 \times (15 - 20) = -60

Labor Efficiency Variance = SP \times (AQ - SQ) = 20 \times (12 - 10) = 40.

The entry to record the labor costs is as follows:

\[
\begin{align*}
\text{Dr. Work-in-process Inventory—Baking} & \quad $200 \\
\text{Dr. Labor Quantity Variance—Baking} & \quad $40 \\
\text{Cr. Labor Price Variance—Baking} & \quad $60 \\
\text{Cr. Wages Payable—Baking} & \quad $180
\end{align*}
\]

*To record standard cost of labor used in production and related labor variances*

We summarize this recording in the following table, which follows the basic structure of the Variance Crown. The left side of the table shows the resources consumed (Wages Payable), while the right side shows the cost applied to Work-In-Process Inventory. The middle column shows the split point, and the two variances are simply the differences between each side column and the center column.

<table>
<thead>
<tr>
<th>Incurred Credit to Wages Payable</th>
<th>Split Point</th>
<th>Applied Debit to WIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$15/hour \times 12 hours = $180</td>
<td>$20/ Hour \times 12 hours = $240</td>
<td>$20/ Hour \times 10 hours = $200</td>
</tr>
</tbody>
</table>

DL Rate Variance \( (AP - SP) \times AQ \) \(-\$60\) Favorable

DL Eff. Variance \( SP(AQ - SQ) \) \$40 Unfavorable

*SQ is shorthand for Standard Quantity Allowed per Units Produced (SQAPUP)*

The footnote at the bottom of the table conveys a key point that many students find confusing:

SQ is shorthand for the less elegant acronym SQAPUP, the Standard Quantity of labor Allowed Per Units actually Produced. In this case, the performance standard allows 5 hours per batch, which translates to 10 hours for the two batches made.

Don’t forget that we apply overhead costs to inventory every day of the year, based on how much the firm produced that day. Every day that passes, we apply a little more overhead. The calculation on the right side of the table above is simply a summary of the overhead applied over the course of the year.

**Standard Costing for Overhead**

We can apply the same central mechanism just applied to direct costs to indirect costs (overhead), but the resulting entries differ in two ways. First, overhead costs include fixed costs as well as variable costs.
Because fixed and variable costs behave so differently, we need to treat them separately. Second, we typically must wait until the end of the period to record overhead variances, because we can’t know the amounts we should have spent on overhead until we know the full activities of the firm (such as the total volume of production).

*Recording and Decomposing Variable Overhead Variances*

Assume that Careful Cookies’s budget specifies that over the course of the year, variable overhead in a given cost pool will total $10,000, to be allocated over the 1,000 machine hours (MHR) allowed to make 1,000 batches of cookies. This budget includes two performance standards: variable overhead should cost $10/MHR (the Predetermined VOH Rate), and each batch of cookies should require 1 MHR. At the end of the period, the actual variable overhead incurred turns out to be $11,000, and it required 800 actual MHRs to make 900 batches of cookies.

Over the course of the year, the firm will make the following entries. First, they will record the incurrence of variable overhead by crediting various resource accounts (cash, accumulated depreciation, wages payable, etc.) and debiting Variable Overhead (similar to a MOH account, but including only variable costs) for the actual cost of the resources:

\[
\begin{align*}
\text{Dr. VOH} & \quad $11,000 \\
\text{Cr. Resources} & \quad $11,000
\end{align*}
\]

To record the incurrence of Variable Overhead

Note that this entry is *exactly the one we would have made* under a normal costing system, because the firm has no way to know whether the VOH resources are being consumed at too fast or slow a rate until they have recorded all actual machine hours.

Second, the firm will apply overhead to Work-in-Process as they create batches of cookies. This entry *differs from normal costing* in that it applies cost according to the amount of the driver that should have been used according to the standard. (Normal costing applies costs according to actual driver use.) In this case, it should have taken 900 MHRs to make 900 batches of cookies, so the entry is:

\[
\begin{align*}
\text{Dr. WIP} & \quad $9,000 \\
\text{Cr. VOH} & \quad $9,000
\end{align*}
\]

To record the application of Variable Overhead at the standard rate and driver quantity

The VOH account now has a $2,000 debit balance, reflecting underapplied overhead. The underapplied overhead arises from two offsetting effects. On the one hand, the firm used 100 fewer hours than allowed by the standard to make 900 batches of cookies. On the other hand, they incurred substantially more variable overhead per MHR than specified in the standard ($11,000 for 800 MHRs, for a cost of $13.75/MHR). We can summarize the breakdown in the same way as we did for direct labor:
Mechanics of Standard Costing

### VOH Variances

<table>
<thead>
<tr>
<th>Incurred Credit Resources/Debit VOH</th>
<th>Split Point</th>
<th>Applied Credit VOH/Debit WIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP x AQ</td>
<td>SP x AQ</td>
<td>SP x SQ*</td>
</tr>
<tr>
<td>$13.75/hour 800 hours $11,000</td>
<td>$10/Hour 800 MHR $8,000</td>
<td>$10/MHR 900 Hours $9,000</td>
</tr>
</tbody>
</table>

DL Spending Variance
(AP – SP)AQ
$3000
Unfavorable

DL Eff. Variance
SP(AQ – SQ)
$3000
Favorable

*SQ is shorthand for Standard Quantity Allowed per Units Produced (SQAPUP)

The firm will record these variances when they close out the MOH account, as follows:

Dr. VOH Spending Variance $1,000
Cr. VOH Efficiency Variance $3,000
Cr. VOH $2,000

To record the VOH variances, which will later be included in the adjustment to CGS

The Spending and Efficiency Variances will be recorded as separate adjustments to Cost of Goods Sold, providing more information than normal costing, which would have combined those adjustments, along with any adjustments for under- or over-applied fixed overhead.

### Recording Fixed Overhead Variances

Now assume that the cost pool just examined is budgeted to include $5,000 of fixed overhead, but the firm actually incurs $5,700. The entries to record the incurrence and application of FOH are similar to those for variable overhead. As the firm receives invoices for rent, insurance, and other fixed costs, they record those at actual cost in a Fixed Overhead account, just as they would under normal costing:

Dr. FOH $5,700
Cr. Resources $5,700

To record the incurrence of Fixed Overhead

As they produce 900 batches of cookies, they apply overhead to inventory according to the amount of driver that should have been used to produce 900 batches of cookies. Given projected FOH of $5,000 over a volume of 1,000 batches at 1 MHR/batch, the FOH rate is $5/batch, so the firm will apply a total of $4,500 of fixed overhead to inventory:

Dr. WIP $4,500
Cr. FOH $4,500

To record the application of Fixed Overhead at the standard rate and driver quantity
Note that this entry is not the same as in normal costing, because it uses the standard quantity of 900 MHrs allowed to produce 900 batches, rather than the 800 MHrs actually used.

At the end of the period, the firm has applied substantially less fixed overhead than it incurred ($4,500 vs. $5,700). Standard costing systems recognize two separate causes for this difference. First, the total amount of fixed cost incurred differs from the amount budgeted ($5,700 vs. $5,000). Second, the total amount of fixed cost applied also differs from the amount budgeted, because the firm produced only 900 batches of cookies instead of the expected 1,000.

Note that the breakdown of FOH is not the same as for VOH, because budgeted and incurred FOH are not the product of a price and quantity, but are simply lump sums. (The amount applied to inventory is indeed the product of production volume and the FOH rate, because absorption costing systems treat fixed overhead as if it were a variable cost.)

To calculate the variances associated with the two separate causes of over- or underapplied FOH, we create a table that is similar to that for VOH, but not identical:

<table>
<thead>
<tr>
<th>FOH Variances</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Incurred</strong></td>
</tr>
<tr>
<td>Credit Resources/Debit FOH</td>
</tr>
<tr>
<td>Incurred FOH</td>
</tr>
<tr>
<td><strong>Split Point</strong></td>
</tr>
<tr>
<td>Budgeted FOH</td>
</tr>
<tr>
<td><strong>Applied</strong></td>
</tr>
<tr>
<td>Credit FOH/Debit WIP</td>
</tr>
<tr>
<td>SP x SQ*</td>
</tr>
<tr>
<td>$5/MHr</td>
</tr>
<tr>
<td>900 Hours</td>
</tr>
<tr>
<td><strong>$5,700</strong></td>
</tr>
<tr>
<td><strong>$5,000</strong></td>
</tr>
<tr>
<td><strong>$4,500</strong></td>
</tr>
<tr>
<td><strong>FOH Budget Variance</strong></td>
</tr>
<tr>
<td>$700</td>
</tr>
<tr>
<td>Unfavorable</td>
</tr>
<tr>
<td><strong>FOH Volume Variance</strong></td>
</tr>
<tr>
<td>$500</td>
</tr>
<tr>
<td>Underapplied</td>
</tr>
</tbody>
</table>

*SQ is shorthand for Standard Quantity Allowed per Units Produced (SQAPUP)

Note that the FOH Volume Variance is listed as “Underapplied” rather than “Unfavorable.” The reason is that, even though the adjustment for the volume variance increases Cost of Goods Sold, producing less than budgeted may well have been the right decision. The last thing Eli Goldratt (author of The Goal) would want an accountant to do is call reducing unnecessary production “unfavorable” or producing more than necessary “favorable”!

The corresponding entry creates two variances that reflect these two distinct causes for underapplied fixed overhead. Both are ultimately reported as adjustments to Cost of Goods Sold.

Dr. FOH Budget Variance  $700
Dr. FOH Volume Variance  $500
Cr. FOH $1,200

To record the FOH variances, which will later be included in the adjustment to CGS
Part IX: The Dark Side
This essay provides frameworks for discussing and reducing illegal and unethical behaviors in the workplace. For illegal behavior, I rely on the framework provided by the Association for Certified Fraud Examiners (ACFE), while for unethical behavior I rely on a framework based on the writings of the philosopher Sissela Bok. There are many other frameworks for analyzing fraud and ethics, but these are useful, coherent, and easily integrated into the subjects we have covered so far.

What Is Occupational Fraud?

**Occupational fraud** is the use of one’s occupation for personal enrichment through the deliberate misuse or misapplication of the employing organization’s resources or assets. By definition, occupational fraud is illegal. The ACFE identifies three primary categories of occupational fraud:

- **Corruption.** Employee’s use of his or her influence in business transactions in a way that violates his or her duty to the employer for the purposes of benefitting him- or herself or someone else. One example would be a *kickback scheme*, in which an employee agrees to purchase from a supplier in exchange for a direct personal payment. Favoritism in hiring is another common form of corruption.
- **Asset Misappropriation.** Schemes in which the perpetrator steals or misuses an organization’s resources. Asset misappropriation can range from the theft of cash to the personal use of company cars and aircraft.
- **Fraudulent Reporting.** Intentional misstatement or omission of material information in the organization’s performance reports. I have generalized the ACFE definition slightly, because the original version referred only to “financial reports.” However, misreporting about nonfinancial performance can also enrich an employee through unwarranted compensation, promotion, or retention (if the employee would have been fired had performance been reported faithfully).

You may wonder why the examples above exclude some of the most egregious forms of illegal business behavior, such as collusion between businesses, willful negligence about consumer safety, money laundering, drug dealing, illegal arms trade, and many activities of organized crime. These are not occupational fraud, in which the employee enriches him or herself at the organizations’ expense; rather, they are perpetrated *on behalf of* the organization. Such crimes fall into the ACFE’s category of **enterprise corruption**, and are outside the scope of this essay. However, misreporting performance to investors (as in the case of Enron, WorldCom, HealthSouth, and other high-profile financial reporting frauds) does constitute occupational fraud, because top executives are employed by their investors.
Determinants of Occupational Fraud

According to the fraud triangle, a model promoted (but not devised) by the ACFE, three factors enable occupational fraud:

- **Pressure**. Employees rarely consider committing fraud unless doing so (without getting caught) will solve a serious problem.
- **Opportunity**. Employees considering fraud rarely act unless they believe they can do so without detection and punishment.
- **Rationalization**. Employees won’t act on an opportunity to commit an undetected fraud unless they can persuade themselves that they are still a moral person.

Preventing Occupational Fraud

The fraud triangle provides a straightforward recipe for reducing the incidence of occupational fraud.

**Identifying and Reducing Pressure**

The first step to preventing fraud is paying careful attention to changes in employees’ personal and financial circumstances. Note that the triangle uses the term “pressure,” rather than “incentive,” to emphasize that people are far more motivated to avoid losses than simply to improve their lot. Few frauds arise just because an employee wants a nicer car or a bigger house. Don’t we all? Frauds far more commonly arise because an employee faces financial or personal difficulties likely to result in a loss of wealth, status, or employment. Effective fraud prevention starts with watching for “red flags” that indicate such losses are possible. Common indicators include substance abuse, divorce, major new expenses (often associated with family members’ health needs), and poor job performance.

Firms can also prevent fraud by not imposing unnecessary pressures in the workplace. Some observers attribute Enron’s culture of fraud to their “rank-and-yank” system of performance evaluation, which required managers to rank the performance of their employees and fire the ones in the lowest categories. While such systems can motivate employees to work harder, they can also motivate them to consider maintaining their present position via less noble means.

**Reducing Opportunity**

Entire courses are devoted to devising control systems that eliminate opportunities for undetected occupational fraud. This essay mentions only one technique, the separation of duties. Organizational transactions typically require more than one action (duty). Decisions must be authorized in principle, and then specifically approved by verifying that the person making the decision has the authority to do so. Assets those decisions affect must be placed in someone’s custody. Someone must examine supporting documents and enter them in the reporting system, and someone must be responsible for the resulting operational and financial performance of the business or unit. One key to preventing fraud opportunities is to separate these four duties. For example, no single person should be able to arrange a
sale of goods and then take immediate custody of the cash received from the customer. No one who is entering data about performance should also be held responsible for that performance.

Limiting Rationalization

To understand how organizations can prevent rationalization, we must understand how people rationalize. In “Moral Disengagement in the Perpetration of Inhumanities” (link), Stanford psychologist Anthony Bandura provides a comprehensive model of the many ways a person can “morally disengage” from reprehensible behavior to maintain a self-image as a good person. He illustrates the process as follows:

Each part of the process suggests its own solution.

- People convince themselves that their conduct is not reprehensible because it is morally justified (they needed to feed their family or provide health care to their ailing parents), because it is not nearly as bad as what others have done, and because it wasn’t “stealing,” it was just “unrecorded borrowing” (euphemistic labeling). Combat these rationalizations in your organization by making sure employees know that fraud is immoral, by making examples of people who perpetrate even minor frauds, and by stamping out tendencies to downplay fraud through euphemisms.
- People convince themselves that they did not actually cause harm; the company is so large, no one is actually hurt. Combat these rationalizations by making sure employees know the true cost of fraud to the company, its clients, the employees’ colleagues, and society.
- People blame the victim: I stole from the company because it treats its workers, customers, or the environment unfairly. Combat these rationalizations by making sure the organization treats workers, customers, and society fairly. Cultivate a positive image among employees, and most
importantly, encourage customers to identify with the organization, rather than seeing it as the “other.” (People are far more willing to blame “others” than themselves or their friends.)

Ethics of Measure Management

A central concern of these essays is that employees will make operational and reporting decisions to manage the measures of their performance, rather than managing the underlying performance those measures are intended to capture. Measure management is rarely illegal, and therefore rarely meets the definition of occupational fraud. To provide insight into when measure management is unethical, I use a framework based on the work of Sissela Bok, philosopher and author of the book *Lying: Moral Choice in Private and Public Life*. We start our analysis by answering two questions:

1. Is an instance of measure management “lying”?
2. If it is a lie, is it harmful?

In this framework, measure management is unethical only if it is a lie and it causes harm.

Bok’s analysis can be viewed as a mix of two philosophical traditional in ethics. Deontological (duty-based) analyses assess the ethics of actions based on whether the act itself violates a duty required of us (such as the duty of honesty). Consequentialist analyses focus instead on whether the action creates harm. The framework here is partly deontological, because Bok believes we have a general duty not to lie. However, consequences still matter; lies that serve good purposes and can sometimes be viewed as ethical.

*Is Measure Management Lying?*

The *Stanford Encyclopedia* includes the following “most commonly accepted” definition of lying (see [http://plato.stanford.edu/entries/lying-definition/](http://plato.stanford.edu/entries/lying-definition/)):

> “I take a lie to be an assertion, the content of which the speaker believes to be false, which is made with the intention to deceive the hearer with respect to that content.” (Williams, 2002. *Truth and Truthfulness: An Essay in Genealogy*. Princeton: Princeton University Press, p. 96)

This definition includes a number of requirements, each of which must be assessed separately.

The first question is whether someone engaging in measure management is making an assertion. Auditors take the view that every line item in a financial statement is an “assertion”—in fact, it includes at least five distinct assertions about the events recorded. Here are the primary management assertions in financial statement line items, which generalize easily to nonfinancial performance reports.

- **Occurrence.** The events reported did occur.
- **Completeness.** All events governed by that line item were recorded.
- **Accuracy.** The numerical measures of all events were appropriate.
Classification. The events were included in the appropriate line item.

Cutoff. The events reported occurred during this accounting period.

A measure of a manager’s performance is therefore an assertion that the reported performance occurred, that there aren’t other aspects of performance that should have been included, and so on.

Is a performance report affected by earnings management “false”? We already know (from the essay “What Counts and What Is Counted”) that a reported measure captures the underlying construct with error. In this sense, every performance report is false.

A more difficult question is whether a report’s falsehood deceives managers. Any good manager (at least one who has read these essays) interprets performance reports with a good deal of skepticism, so although any performance report with measurement error can be viewed as false, it won’t deceive good managers. Moreover, good managers also interpret performance reports with a good deal of cynicism—skepticism born of the belief that reporters bias performance measures to their own advantage. To the extent that managers expect some degree of opportunistic reporting and real measure management, one can argue that such behaviors do not deceive the manager, and in fact avoiding these behaviors could deceive the manager even more (because their cynicism would be unwarranted).

A report’s deceptiveness is ultimately determined by whether the measure management is consistent with the reader’s expectations. For example, a school superintendent might expect a teacher to “teach to the test” that measures student performance, and a hospital director might expect a surgery unit to transfer high-risk patients to facilities with less onerous reporting requirements. However, few would expect the teacher to change students’ test answers, or a surgery unit to state that patients survived an operation when they did not.

My own analysis (not Bok’s) is that we can best assess the deceptiveness of a report by asking whether those with discretion over operations and reporting followed the rules of engagement for measure management, which set the expectations for allowable and unallowable behavior. For external financial reporting, reporting and auditing standards set most of the rules of engagement. Investors would be deceived if a manager reported sales commissions as assets (a classification assertion), or reported a January sale in December (a cutoff assertion), because these actions are clearly forbidden. However, investors would not be deceived if the investors discounted sales in December to increase income before the year ended, because such behavior is not only permitted by standards, it is common.

Internal reporting is governed not by regulatory standards, but by top executives. Many firms establish strict codes of conduct that specify what operational and reporting choices are acceptable. Teachers are typically given permission to “teach to the test,” but changing student test answers clearly violates established policies and is therefore deceptive. If a hospital had a stated policy not to make treatment decisions based on their influence on error reports, doing so would be deceptive (even though it might not be at another hospital that lacked such a policy).

Some additional points about rules of engagement are worth emphasizing:
1. Firms have broad powers to set their rules of engagement. They should do so with an eye to maximizing the usefulness of their reporting outcomes—are their reports as useful as they can be? More relaxed rules of engagement may be useful if proxy measures are inherently noisy, because managers may be able to communicate more effectively if they have more discretion to 

2. While firms can set their own rules of engagement, in practice most managers seem to share the view that many forms of operational measure management lie within the rules of engagement, but that opportunistic reporting is within the rules only if it involves difficult judgments (usually about classification or measurement) that are hard to verify objectively. Almost no one believes that altering data lies within the rules of engagement.

3. The primary ethical challenges lie in choices that push the bounds of measure management—they don’t clearly violate the rules of engagement, but they are rare or perhaps novel. For example, an innovative sales group might devise a new type of contract that allows them to report high revenue in this period according to the rules the organization uses to measure revenue, even though services need to be provided in future periods.

If It Is a Lie, Is It Still Permitted?

Sissela Bok’s analysis of acceptable lies rests on her *Principle of Veracity*, which asserts that there is a strong moral presumption against lying. However, that doesn’t mean you can never lie. Some lies can be praiseworthy (“You look great in that outfit.”) If you are considering lying, Bok recommends asking whether your choice and the reasons for it would impress other reasonable people.

People are likely to be persuaded by a mix of deontological (duty-based) and consequentialist elements. For example, many people oppose lying if you could have accomplished your goals more honestly (because you have a duty to be honest). Could you have avoided distorting operational or reporting decisions by persuading your supervisor to pay you a bonus despite misleadingly poor reported performance? If so, Bok would clearly argue for that approach.

People are also more opposed to your lying if it has bad outcomes for the organization or others. Will measure management outside the rules of engagement cause a manager to part with a bigger bonus (to you!) even though it is unwarranted? Will the action mislead the manager into severing ties with a customer, supplier, or employee (not you!). These bad outcomes suggest that lying is not warranted. On the other hand, perhaps measure management (or sandbagging during the budget process) helps the organization by ensuring that they give much-needed resources to your department.

Finally, Bok emphasizes that you must consider a less-obvious consequence of lying: “the effects of your lie on the *general practice of veracity* itself.” Lies, like pollution, carry externalities when they weaken the rules of engagement. Even if a lie does not impose a direct cost on the people who are relying on your statements, it may impose an indirect cost by making it easier for others to believe that “this is how the game is played.” It is hard to see how you can be an ethical person if you lie without considering these costs.
Behavioral Biases and Management Reporting

Small Brains and Strong Emotions

A primary goal of accounting is to provide information useful for decision-making. However, business decisions are not made by cold, intellectually powerful “Spock-like” computers, but by people with small brains and strong emotions, which lead them to place too little weight on some types of information and too much on others. To accommodate these limitations, accountants and managers must often choose what might seem like an inferior reporting system.

Fortunately for accountants, decades of psychological research allow us to predict what types of poor decisions to expect from users of our reporting systems. Many poor decisions are not simply random mistakes (a form of noise), but are actually predictable, systematic errors (a form of bias). We make some of these errors because our brains are simply too small and inefficient to make the right decision every time. Instead, we follow simple rules of thumb, called heuristics, which are often correct but predictably lead us astray in certain settings. These errors are called cold biases, because they result simply from lack of computing power. We make other errors because we have built-in psychological tendencies to protect ourselves from unpleasantness. Such errors are called hot biases, and these can be far more difficult to identify and overcome.

What follows is a litany of cold and hot biases, along with advice on how to recognize and overcome them, both as a decision-maker and as a designer of reporting systems.

Cold Biases

The Availability Bias

Are celebrities more likely to get divorced? Does wearing your lucky socks help your favorite team win? Is your assistant manager likely to have errors in his PowerPoint slides for the most important meeting of the year? Whether these claims are actually true depends on a large body of data. We would need to know how often the event in question occurs (divorces, wins, errors) for the entire population (of celebrities and regular folks, with and without lucky socks, and for important and unimportant meetings). But we rarely have access to that type of comprehensive analysis. Instead, whether we think these claims are true depends on how easily we can recall examples of such events. This “availability heuristic” is generally a reasonable rule of thumb. If there really is a strong tendency for something to happen a certain way, we see many occurrences of it during our lives. Memory researchers have shown that each time the event is stored in our memory, it leaves what is called a “memory trace.” The more
Behavioral Biases and Management Reporting

memory traces exist for a certain type of event, the easier it is for us to recall when we are asked whether something tends to happen. Such events are said to be more available.

The availability bias arises when memories are more available, even when they aren’t actually more common. It is easy to think of many celebrities who have divorced, but that is because their personal lives are highly publicized and become common topics of conversation. It is easy to think of times our favorite team won when we wore our lucky socks (or lost when we didn’t), because such events are emotionally laden and surprising (because most of us know that our socks can’t really affect events broadcast on television hundreds of miles away). These events leave more or deeper traces, causing us to overestimate how commonly they occur. It is much harder to think of ordinary people who divorced, the times our team won or lost when we weren’t thinking about socks because we couldn’t watch the game, and the times there were errors at unimportant meetings or error-free important meetings.

- Mitigate your own availability biases by paying careful attention to whether the event you think is common is more widely publicized or more emotionally salient that the alternative event. If it is, seek out broader data.
- Help your company by identifying settings in which availability biases will affect your colleagues. Conduct a large-scale analysis of the data your reporting system collects, and make sure to publicize the results. Let everyone know whether (for example) costly errors are more likely when people are tired.
- Take advantage of the availability bias by publicizing events you want people to remember, or making them more emotionally salient.

The Halo Effect

“Give a man a reputation as an early riser and he can sleep till noon.”

—Benjamin Franklin

Modern psychological research corroborates Benjamin Franklin’s intuition that performance evaluations are heavily influenced by what we already know about a person. In a famous experiment from the 1960s, two groups of subjects were shown a picture of a well-groomed, strong-looking, dark-haired man looking confidently at the camera, and asked to rate how physically handsome he was. One group rated him as much more attractive than the other. Why? The only difference was that the group rating him highly was told he was a movie star. The other group was told that he was a serial killer. Remember, the subjects were not asked to rate the “inner beauty” of the man in the picture. But like most people, the subjects in the experiment couldn’t separate two dimensions of evaluation: how the person looks and what they do for a “profession.” This type of bias is known as the halo effect.

Like most cold biases, the halo effect is the predictable downside of a useful heuristic. Managers who perform well in one year tend to perform well the next, and managers who are good at one task tend to be good at another. (And despite what you think of celebrities, people who are attractive also tend to be intelligent.) But these are only tendencies. A manager may have encountered personal problems since
last year that distracted her from her duties. One task may be less well suited to the manager’s strengths than another. These are cases in which the halo effect can lead us astray.

- Protect yourself from the halo effect by ignoring information that isn’t directly relevant to the performance you are evaluating. When evaluating this year’s performance, don’t look at last year’s performance, or focus on changes from last year to this year. Evaluate performance on one task at a time.
- Protect your organization from the halo effect by limiting information that is likely to distort performance evaluations. Take a cue from orchestral auditions: players are seated behind a screen so that the judges cannot be swayed by any information but the quality of the music they hear.

**Functional Fixation**

Karl Dunker conducted an experiment in 1945 that showed how hard it is to see new uses for familiar objects. Dunker gave subjects a box of thumbtacks and a candle, and asked them to affix the candle to the wall. The solution eluded most of the subjects: use the box itself as a candleholder, tacking it to the wall and tacking the candle onto it. Subjects were said to be “fixated” on the original function of the box, leading Dunker’s experiment to be famous as the first example of the *functional fixation* bias.

Functional fixation is a serious problem in managerial reporting, because people become accustomed to using reported measures in a certain way. Changes in the reporting system, the way the measure is calculated, or environmental circumstances may change the function the measures should perform, but people get functionally fixated and don’t change their approach as much as they should. For example, for many years US firms did not need to report an expense for the stock options they gave to employees. The options were clearly valuable (as Warren Buffet famously said, “If they aren’t valuable, give them to me!”), so the FASB altered GAAP to force firms to report this expense. The change caused a dramatic reduction in earnings, but had no effect on the true value created by a firm in the reporting period—it affected only the meaning of the earnings number that attempted to measure that value. However, managers of high-tech companies that relied heavily on option-based compensation worried that investors would treat the new earnings numbers just as they did the old ones.

Functional fixation has also distorted the interpretation of polling data businesses and politicians use. For decades, pollsters have randomly called phone numbers in an attempt to get representative samples of customers and voters. Over the last five years, however, vast numbers of young and low-income Americans have given up their landline phones, relying entirely on cellphones, which have been much harder for pollsters to reach. Only recently have pollsters begun to adjust for the shifting demographics of their phone-based polls, and many still attempt to interpret their numbers as they always have.

The advice for avoiding functional fixation is simple: make sure you understand how numbers are calculated, so that you can modify your interpretations as needed.
Hot Biases

*Motivated Reasoning*

“It is hard to get a man to understand something when his salary depends upon his not understanding it.”

—Upton Sinclair, 1935

Some of the most pernicious hot biases are those that lead us to the conclusions that we prefer. These biases go by many names, including “overconfidence,” “wishful thinking,” “affirmation bias,” and “confirmation bias.” But I prefer the term “motivated reasoning” because it so plainly expresses the basic problem—we distort the way we reason because we are motivated to reach a certain conclusion. *Motivated reasoning* pervades all aspects of our reasoning:

- We seek out information that confirms what we want to believe, and avoid information that counters it.
- We place more weight on information that confirms what we want to believe than on information that counters it.
- Once our reasoning leads us to preferred conclusions, we stop reasoning, but if it leads us to unpleasant conclusions, we seek further information and reevaluate the information we already have.

Like most hot biases, motivated reasoning is extremely difficult to prevent, partly because human psychology seems to have evolved to encourage it. But here are some useful strategies:

1. Be aware of the conclusions you (or others in your organization) prefer to believe Force yourself to seek out contrary information, and resist the urge to keep searching when the evidence is against your preferred conclusion. Delegate information-collection and analysis to those who do not have a strong stake in the conclusion.
2. Limit the number of performance measures you report, to limit opportunities for employees to convince themselves they are performing better than they are. If you evaluate performance on 27 different measures, the people being evaluated will focus on the measures that tell the most favorable story, and will be surprised if they find out that their employers care more about the others. By limiting the number of performance measures, you limit opportunities for motivated reasoning.
3. Separate evaluations of outcomes from evaluations of what a person “really is.” Motivated reasoning, like other hot biases, is triggered by our strong desire to protect our identities: we want to continue believing what we currently believe about ourselves. Most managers identify themselves as moral, popular, successful, and effective people. Create a culture that recognizes that bad outcomes don’t imply bad character.
The Fundamental Attribution Error

Attributions of credit and blame are notoriously difficult, even without the threat of hot biases, because outcomes almost always reflect a mix of both the actions of the people involved and the environments in which they act. The fundamental attribution error is that we ascribe other people’s success or failure too much to their own contributions and too little to environmental factors. We are far too quick to attribute some else’s failure to a lack of effort or intelligence, rather than to a difficult task, inadequate resources, or bad luck. Similarly, we think of someone who lies as a “liar” with poor moral character, rather than as someone who was perhaps placed in a difficult position.

The following graph shows just how mistaken we usually are when we attribute bad actions to bad character. The graph below depicts the famous experiments on obedience by Stanley Milgram (as re-reported by Anthony Bandura). Subjects in the experiments were asked to administer electric shocks to a person they thought was another subject. (In fact, it was a confederate of the experimenter who only pretended that the shocks were painful.) When asked to administer the shocks by an authority figure (a man in a white lab coat) and surrounded by peers who complied with the request, over 90% of the subjects obeyed the command. When the authority figure moved to a distant location and conveyed commands over the phone, only 20% obeyed, and when two authority figures issued conflicting commands, no one obeyed. The stark differences show that “moral character” had little effect on subject’s behavior—environmental factors were far more influential.

Combination Biases

Many flawed judgments and decisions arise from a combination of cold and hot biases. There are many such biases; this essay’s conclusion focuses on two that are particularly relevant to the workplace.

The Planning Fallacy

How long will it take you to finish your next project for school? When will you complete your next major report for work? If you are like most people, you significantly and consistently underestimate how long those and other complex tasks will take. The planning fallacy—underestimating the time required to complete complex tasks—arises from a combination of cold and hot biases. Our brains are not very good at remembering all of the commitments we have made, or anticipating surprises that will interfere with a task that may take weeks or months to complete. We also want to believe that we have enough time to complete tasks we have already committed to.

The best way to avoid the planning fallacy is to take an outside view. Suppose a colleague tells you his complex task will take three weeks to complete. Will you believe him? Most people don’t, because they know that most people underestimate how long it takes to complete a complex task. Just remember that you are “most people” too. Thinking about the specific task at hand and how you are going to schedule your time from day to day leads you to discount the impact of interruptions and surprises. Instead, take an outside view—think about how long someone else would predict it will take you, and rely on that as your guide.

Escalation Biases (and the Sunk Cost Fallacy)

Assume you have devoted a good deal of time, money, or other resources to a project. When it starts to fail, will you throw good money after bad, even though the incremental benefit does not justify the incremental cost? Many people do, falling prey to the escalation bias at substantial cost to their organizations. There are some rational reasons to do so—even if additional resources might not survive a cost-benefit analysis for your organization, it might do so for your own career by delaying people’s recognition that your original investment was unwise. However, a mix of cold and hot biases also leads people to escalate investments when they shouldn’t. Our small brains make it difficult to distinguish costs that have already been committed and can’t be reclaimed (sunk costs) from costs that we can still avoid, an error called the sunk cost fallacy. And our desire to see ourselves in a favorable light makes it all too easy to convince ourselves that a losing project is still worth pursuing.

Good reporting systems can help mitigate escalation biases and sunk cost fallacies by designing reports to distinguish clearly between costs already incurred from those over which the organization still has control.
GAAP: Generally Accepted Accounting Principles. Set of accounting standards as established by FASB and/or IASB. No GAAP for managerial accounting because there is no conflict of interest between the reporter and user, and no need for comparability across firms.

Perquisite (Perk): Benefits incidental to employment, such as visibility, pleasant employees, a corner office or rights to use the company jet.

Rent Seeking: Income (including perquisites) extracted from existing capital, including human capital, decision-making rights, expertise, or information.

Bureaucratic Systems: Systems that apply standard rules every time a set of officially-sanctioned conditions apply.

Legacy Systems: Systems still in use that were developed using outdated technology or to solve problems that may no longer be relevant.

Chesterton’s Fence: Shorthand for Chesterton's advice not to tear down a fence (by analogy, a legacy system) before knowing its original purpose.

Just World Hypothesis: A psychological tendency to accept as just bad outcomes that we can't control.

Scientific Management: The practice of applying scientific principles to production and organizational processes.

Taylorism: Synonym for scientific management, in honor of its first public proponent, Frederick Winslow Taylor (book published in 1911).

Human Relations Critique: Scientific management's failure to recognize worker's psychological needs and responses.

Organizational Critique: Scientific management's failure to recognize the imperfections of complex reporting structures and top-down decision-making.

Strategic Critique: Scientific management's failure to recognize that people respond strategically to being measured, in ways that make measurements less useful than in physical sciences.
Glossary

**Accounting System**: Reports financial information using double-entry bookkeeping. One of the five subsystems of management reporting systems

**Budgeting System**: Formalizes an organization's financial and operational plans. One of the five subsystems of management reporting systems

**Performance Reporting System**: Evaluates and improves decision-making, often by comparing results from the accounting system to the budget created by the budgeting system. One of the five subsystems of management reporting systems

**Incentive System**: Links management reports (especially performance reports) to compensation. One of the five subsystems of management reporting systems

**Internal Control System**: Ensures that employees comply with organizational policies, and that data from the other systems are accurate. One of the five subsystems of management reporting systems

**Financial Reporting Standards**: Regulations from FASB and/or IASB that govern external financial reporting. Necessary due to conflicts of interest and the need for comparability across firms.

**Attention-directing**: The ability of a reporting system to induce people to pay attention to, and talk about, one issue instead of another. Think flashing lights and sirens.

**Decision-facilitating**: The ability of a reporting system to help people achieve their own objectives. It gives people the right information to make a decision. Think guide dog. Note the emphasis on "their own" objectives, which distinguishes facilitation from influence.

**Decision-influencing**: The ability of a reporting system to change another person's decision because they are attempting to garner rewards or avoid punishment by superiors. Think 'carrot and stick'. Note the emphasis on the control of OTHER PEOPLE's behavior, which distinguishes influence from facilitation.

**Coordination-facilitating**: The ability of a reporting system to allow people and units to coordinate their decisions, typically through a budgeting and standard-setting process. Think of a conductor's baton, which tells everyone in an orchestra when and how to play together.

**Balanced scorecard**: A representation of a strategic plan as a set of objectives, grouped into categories, and linked in cause-effect relationships.

**Financial Perspective**: One of the four categories of the balanced scorecard. Asks the question, "To succeed financially, how should we appear to our shareholders?"

**Customer Perspective**: One of the four categories of the balanced scorecard. Asks the question, "To achieve our vision, how should we appear to our customers?"

**Internal Process Perspective**: One of the four categories of the balanced scorecard. Asks the question, "To satisfy our shareholders and customers, at what business processes must we excel?"
Glossary

**Learning & Growth Perspective**: One of the four categories of the balanced scorecard. Asks the question, "To achieve our vision, how will we sustain our ability to change and improve?"

**Objective**: The dimension of performance that the manager wishes to improve. Often stated in general terms.

**Measure**: A source of data that can be used to assess performance on an objective.

**Target**: A level of the measure against which performance will be compared to determine success. 90% satisfaction rate; less than 1% defect rate; etc.

**Initiative**: A plan or strategy for achieving a target level of performance.

**Four Food Group Perspective**: The original view of the BSC: a balanced diet of information requires financial, customer, internal process and learning & growth data. Balanced Scorecard 1.0

**Balanced Scorecard 2.0**: Uses the BSC to represent the firm's strategic plan. Each perspective is a link in a causal chain.

**Strategy Map**: Shows the causal links among the objectives of the BSC.

**Leading measures**: A variable whose change is associated with a later change in another (lagging) variable. On-time delivery leads improvements in customer satisfaction.

**Lagging measures**: A variable whose change is associated with a previous change in another (leading) variable. On-time delivery leads improvements in customer satisfaction.

**Organizational Units**: A separately-identifiable piece of a company. E.g. departments, divisions, etc. Multiple organizational units make it difficult to implement only a single balanced scorecard.

**Cascade**: Process by which a BSC works in a large organization. Includes three major practices: everyone understands the top-level, each unit has its own BSC, and scorecards are integrated across units.

**Outputs**: Measures that quantify the amount of a service provided, such as miles of road repaired.

**Outcomes**: Measures that gauge the accomplishments or results that occur (at least partially) because of the services provided—for example, the quality of roads as rated by residents.

**Outcome-Based Contracts**: Contracts that specify payments based on outcomes, rather than outputs; very popular with government contract work.

**Pay for Performance**: Contractual or non-contractual links between measurable variables and rewards or punishments meted out on the agent by the principal. Can be explicit or implicit.
Explicit Incentive Compensation: Specifies a contractual relationship between performance measures and current compensation. If the firm earns at least $5,000,000 the executive will receive a $100,000 bonus.

Implicit Incentive Compensation: Links promotions, raises, and non-cash honors to subjective evaluation that is informed by, but not contractually tied to, performance measures. More common than explicit incentive compensation.

Motivation: The desire to accomplish an objective. Also one key effect of incentive compensation. Incentives are NOT motivation but rather may give rise to it.

Communication: The ability of incentive compensation plans to communicate the principal's desires to the agent.

Screening: The ability of incentive compensation to attract agents of particular types. Often an unintended effect.

Risk-sharing: The ability of incentive compensation plans to spread risk between the principal and the agent. Often an unintended effect.

Input: A choice by an agent, such as the decision to work harder or smarter.

Working Harder/Working Smarter: Working harder means providing a greater quantity of effort; working smarter means directing your effort toward more appropriate ends. Effort vs. allocation.

Output: Something that results in part by an agent's input and is valued (either positively or negatively) by the principal.

Outcome: Results valued by a principal that occur due to output. The distinction between output and outcome is important in service industries. Hours of training is an output; learning is an outcome.

Controllability Principle: Workers should be compensated on the basis of measures over which they exercise a high degree of control. Don't pay shirt-makers based on stock price.

Risk aversion: A preference of less risk over more risk.

Compensating differential: An increase in pay to compensate a worker for taking on risk they wouldn't face in an otherwise identical job that imposes less risk. Must compensate for financial, health, and safety risks.

Self-select: A decision by a prospective agent to take or not take a job based on their inside information about their own traits and their knowledge of the job. You self-selected not to play professional basketball.

Information Asymmetry: When one party knows more than another party. Often causes problems in contracting and employment.
Effort aversion: A preference for exerting less effort to exerting more effort.

Principal: A person who assigns another (the agent) to undertake a task. "The Boss" or "The Owner"

Agent: One who undertakes a task on behalf of another "The worker"

Residual Claimant: The person who gets the remaining profit after all other costs have been paid. The principal is the residual claimant because she keeps the value created by the agent's performance

Incentive Intensity: The sensitivity of pay to a performance measure. The slope of the pay-performance function

Moral hazard: A circumstance in which an agent is shielded from the full impact of his decisions. Agents almost always face moral hazard because the principal cannot perfectly distinguish luck from effort.

Participation Constraint: The recognition that the principal must be sufficiently generous to induce the agent to self-select to take the offered job.

Incentive Compatibility Constraint: The recognition that the agent will make decisions that are in his own interest, not in the interest of the principal

Disagreement-based Principal-Agent Models: Models in which the principal's concern is that the agent has different preferences over output, and therefore may disagree with the principal's policy. Incentives in these models are used not to motivate effort, but rather to alter the agent's performance goals.

Intrinsic Motivation: A personal desire to meet an objective because doing so is psychologically rewarding. Children enjoy reading because stories are entertaining.

Extrinsic Motivation: A desire to accomplish an objective because it will result in non-psychological rewards, typically financial compensation. Kids will do things they don't enjoy if you reward them with pizza.

Crowding-Out Effect: The tendency for extrinsic incentives to reduce and replace pre-existing intrinsic motivation. Promising pizza to encourage kids to read books will create "Fat Kids Who Hate to Read"

Forms: The true reality as explained in Plato's Cave Allegory

Shadows: Our incomplete representation of reality as explained in Plato's Cave Allegory.

Skepticism: The belief that we cannot trust our senses (proxy measures) to give reliable information about underlying reality (constructs).

Proxy measure: A variable used as a measure of an underlying construct. What present-day scientists would call Plato's shadows
**Glossary**

**Underlying construct**: An idealized variable of interest that cannot be observed, but is thought to have a causal or non-causal association with other constructs. "Forms" in Plato’s cave. We see only their shadows (proxies) on the wall.

**Cause**: A cause is the construct that induces a change in another construct (the effect). Hard work is the cause, learning is the effect.

**Effect**: An effect is a construct that changes due to a change in another construct (the cause). Hard work is the cause, learning is the effect.

**Omitted Variables**: An omitted variable is one that might affect the proxy measure for the final effect, but which we do not include as part of our causal model (i.e. not a "cause"). Intelligence and stress are omitted variables that affect grades.

**Measurement error**: The difference between the true value of a construct and the value of its proxy measure. Consists of noise and bias.

**Noise**: The random or unpredictable component of measurement error between the proxy and construct. A measure based on a small sample is noisy. So is a measure influenced by unknown omitted variables.

**Bias**: The predictable directional component of measurement error between a proxy and construct. People tend to think they work harder than they do—an upward bias.

**Precision**: The absence of noise in a proxy measure.

**Mediating variable**: A variable that lies in between a cause and an effect. On-time delivery mediates machine uptime and customer satisfaction.

**Moderating Variable**: A variable that increases or reduces the effect of another cause. Talent moderates the effect of practice on performance.

**Campbell’s Law**: The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.

**Measure management**: Taking actions that introduce a desired bias into a proxy measure (relative to the underlying construct it is intended to capture).

**Law of Measure Management**: Elaboration of Campbell’s law, recognizing that measure management arises when proxy measures capture constructs with error, people know how they are being evaluated, and they have discretion over how raw data is generated or transformed.

**Real/Operational Measure Management**: Measure management achieved by altering operational decisions that affect raw data used to generate performance measures.
Opportunistic Reporting: Measure management achieved by altering how raw data are transformed into performance measures.


Entity: Any social organization (e.g. individual, family, firm, government) whose wealth is being tracked.

Assets: Probable future economic benefits obtained or controlled by a particular entity as a result of past transactions or events.

Liabilities: Probable future sacrifices of economic benefits arising from present obligations of a particular entity to transfer assets or provide services to other entities in the future as a result of past transactions or events.

Equity or Net Assets: The residual interest in the assets of an entity that remains after deducting its liabilities. Assets = Liabilities + Equity.

Contributions: Increases in equity of an entity resulting from transfers to it from other entities of something valuable to obtain or increase ownership interests (or equity) in it.

Distributions: Decreases in equity of an entity resulting from transferring assets, rendering services, or incurring liabilities by the enterprise to owners. Distributions to owners decrease ownership interest (or equity) in an enterprise.

Comprehensive Income: The change in equity of an entity during a period from transactions and other events and circumstances from non owner sources. It includes all changes in equity during a period except those resulting from investments by owners and distributions to owners.

Stock variables: Variables that describe the state of an entity at a given moment in time. E.g. Assets, liabilities, equity.

Flow variables: Variables that describe changes from one moment in time to another (usually a month, quarter, or year) E.g. Revenues, expenses, capital changes.

Revenue: Inflows or other enhancements of assets of an entity or settlements of its liabilities (or a combination of both) from delivering or producing goods, rendering services, or other activities that constitute the entity’s ongoing major or central operations.

Gains: Increases in equity (net assets) from peripheral or incidental transactions of an entity and from all other transactions and other events and circumstances affecting the entity except those that result from revenues or investments by owners.
Glossary

**Expense**: Outflows or other using up of assets or incurrences of liabilities (or a combination of both) from delivering or producing goods, rendering services, or carrying out other activities that constitute the entity's ongoing major or central operations.

**Losses**: Decreases in equity (net assets) from peripheral or incidental transactions of an entity and from all other transactions and other events and circumstances affecting the entity except those that result from expenses or distributions to owners.

**Inventory**: Assets that are held for sale to customers or for conversion into an asset that is held for sale to customers

**Property, Plant, and Equipment**: Physical assets that are vital to ongoing operations, but are not held for sale to customers, or converted into an asset held for sale to customers.

**Cost of goods sold**: The expense incurred because the entity has parted with goods or services conveyed to customers as part of ongoing operations.

**Selling, general, and administrative** (SGA): Expenses incurred to maintain ongoing operations other than Cost of Goods Sold.

**Product cost**: Costs that pass through FG into CGS.

**Period costs**: Costs that pass into SG&A.

**Transaction Entry**: Entry made to reflect a transaction. One of the three reasons to make a journal entry.

**Adjusting Entry**: Entry made to adjust an account balance to reflect a change in circumstances other than a transaction. One of the three reasons to make a journal entry.

**Closing Entry**: Entry made to close out an account that needs to have a zero balance before the next reporting period begins. One of the three reasons to make a journal entry.

**Predetermined overhead rate**: The rate used to apply manufacturing overhead to work-in-process inventory, calculated as: estimated overhead pool amount ÷ estimated activity base

**Overhead pool**: Shared resources consumed by a subset of productive activities. Typically associated with a single driver or department.

**Overhead driver**: A measurable variable used to allocate overhead costs to individual cost objects. Must be measurable for each cost object.

**Absorption (Full) costing**: A method of product costing in which both variable and fixed manufacturing overhead are included in product costs. It derives its name from the notion that the products "absorb" all costs, including fixed costs. Also called full costing. Soupcon, Bandway, Hematest and VirtuCall all have absorption costing systems. In Soupcon, the new product affected the margins of the old product
because they used an absorption costing system. In VirtuCall, absorption costing is what made costs rise for other contracts when one contract was cancelled.

Activity base: A measure of an organization's activity that is used as a basis for specifying cost behavior. Activity base is also called "denominator volume"

Normal costing: A product costing system that applies overhead based on a predetermined overhead rate and actual driver use, and measures direct costs based on actual use. Used by HemaTest, VirtuCall and all examples and exercises from Hilton Chapter 3.

Overapplied/Underapplied overhead: The amount by which the period's applied manufacturing overhead exceeds/falls short of incurred manufacturing overhead.

Adjusted cost of goods sold: The cost of goods sold after all adjusting entries (e.g. closing out MOH account) have been made. Underapplied overhead will cause adjusted CGS to be higher than unadjusted CGS. Overapplied MOH will cause it to be lower

Unadjusted cost of goods sold: Cost of goods sold before adjusting entries. This amount is the sum of the individual product reported costs and does not include the adjustment for over or underapplied MOH.

Cost: A measure of the resources consumed or given up to achieve a particular purpose.

Expenditure: A sacrifice of cash or a cash equivalent.

Resource Spending: Expenditures to provide resources that can be consumed.

Resource Consumption: The use of a resource already owned.

Out-of-pocket cost: Expenditures.

Prime costs: Conceptually, direct material cost + direct labor cost. Operationally, costs added to those accounts.

Capacity Cost: The consumption of a resource that, once purchased, allows a limited range of production volume. Think of a machine that can be used 24 hours a day, seven days a week.

Opportunity cost: The potential benefit given up when the choice of one action precludes selection of a different one.

Proprietary cost: A sacrifice of competitive advantage. Often the largest cost of disclosing information.

Cost Assignment/Application: When costs are transferred from one account to another. For example, when raw material is applied to inventory
**Glossary**

**Cost Object**: The account to which a cost is assigned, applied or allocated. Common cost objects are jobs, departments, clients, products, product lines.

**Job**: A specific product or service. A separate cost object

**Job Shop**: A business that keeps track of costs on a specific product or service level (i.e. job level). Often done with specialty services, repairs, construction, etc. Never done for mass-production

**Job Order Costing System**: A costing system that allocates costs to items produced or services provided for each specific customer demand (job).

**Direct Materials**: Material costs that can be tracked by the individual job

**Direct Labor**: Labor costs that can be tracked by the individual job

**Shared Resource**: A resource that could be used to create two or more outputs, but need not produce both. A machine that can be used to produce EITHER product A OR product B is a shared resource, but not a joint cost.

**Indirect cost / Overhead**: Product costs that cannot be directly applied to a particular good or service.

**Process Shop**: A shop that does exactly the same thing over and over. Does not need to track costs by the individual unit because every unit is the same.

**Processing/Conversion Cost**: Direct labor plus overhead, which can be combined when every item in a department is similar enough to receive the same amount of both costs. Raw materials do not go into conversion costs.

**Mass Customization**: A production process in which every process is done repeatedly (as in a process shop) but output is customized (as in a job shop) by altering which processes are included in a particular product or service.

**Operations Costing**: A form of costing appropriate in mass customization, in which the costs of each process are calculated as in process costing, and then added up for each unit of output according to the processes it went through.

**Product Differentiation**: Features that make a product unlike competing products, allowing the seller to raise prices without eliminating demand

**Commodity Product**: A product with no product differentiation, so that raising prices above that of competitors results in no demand at all.

**Overhead Allocation**: The act of assigning overhead to individual products using one or more overhead drivers.
**Burden rate**: A synonym for overhead cost and indirect cost. Called "burden rate" to emphasize the decision-influencing function of overhead application. "If you want to kill something, tax it!"

**Hidden difference**: A difference in circumstances that could be used to calculate very different overhead rates for a single driver, but isn't being used to do so. Labor hours drive overhead in a consulting firm, but Partner labor drives far more overhead per hour than does Associate labor. A firm that calculates a single overhead rate for labor is ignoring a hidden difference (worker type).

**Hidden driver**: A variable that is positively associated with overhead cost in a cost pool, but is not used to allocate overhead even though it could be. Machine hours were a hidden driver in Hematest's original system.

**Two-stage costing system**: A two-step procedure for assigning overhead costs to products or services produced. In the first stage, all production costs are assigned to overhead pools. In the second stage, costs in the pools are applied to products and services based on the overhead driver associated with that pool. HemaTest!

**Peanut-butter principle**: Overhead costs are like peanut butter. Changing overhead pools and drivers will spread it around differently across cost objects, but won't change the total amount to be spread.

**Balancing rule**: The overhead allocated to a cost object won't change if the drivers used by the object are in exact proportion to the activity bases used to allocate those drivers.

**Cross-subsidize**: From the peanut-butter principle, a bad allocation system won't make costs too high or too low across the board. Instead, some jobs will be priced too low at the expense of other jobs that are priced too high.

**Winner's Curse**: The tendency of the winner of a competitive market to lose money, when the value of a purchase (or the cost of production) is hard to estimate. When a company sets selling prices based on poorly estimated costs, they are more likely to sell the goods when their cost is low and not sell the goods when their cost is high.

**Adverse Selection**: Winners of bids are selected for a reason that is not in their best interest. Winners of auctions often win because they bid too aggressively.

**Shade Bids**: Bidding below your best estimate for value or above your best estimate of cost in order to avoid succumbing to the winner's curse. Prospective buyers should shade their bids downward in order to secure a profit.

**Common Value Auction**: An auction where the true value is the same for all potential bidders. In a common value auction, the winner of a bid has no reason to think that they have a competitive advantage.

**Common Cost Auction**: An auction in which no seller has a significant cost advantage over any other. The seller with the lowest estimate in a common cost auction probably underestimated their costs.
Cost-Volume-Profit (CVP) Model: A generalization of break-even analysis based on the formulas $C = F + vQ$ and $R = pQ$.

Fixed cost: A cost that does not change in total as activity changes.

Activity Volume: Amount of activity used to produce a given level of product or service. Usually the volume of a product or service provided

Variable cost: A cost that changes in total direct proportion to a change in an organization's activity volume.

Marginal cost: Another name for Variable cost. Emphasizes that this is the cost to produce one more unit.

Contribution Margin: Sales price minus the unit variable cost. This is the amount that each unit "contributes" to covering fixed costs.

Cost Structure: The relative proportions of an organization's fixed and variable costs.

Incremental/Relevant cost: The increase in cost from one alternative to another.

Operational leverage: Fixed costs divided by total cost. Measure of the sensitivity of profit to changes in volume. A firm with high fixed costs and low variable costs has high operational leverage. Therefore, profits will be very sensitive to volume.

Operating leverage factor: A measure of the operating leverage at a particular sales volume. Computed by dividing an organization's total contribution margin by its net income. $\% \text{ Change in Profit} = \text{OLF} \times \% \text{ Change in Sales}$

Full (Absorption) Costing: A system that includes an allocation of fixed costs, along with variable costs, in the reported cost of a product or service.

Marginal/incremental analysis: An analysis that considers the effect of a small increase or decrease in activity.

Variable costing: A method of product costing in which only variable costs are treated as product costs; all fixed costs are treated as period costs. Also called direct costing. VirtuCall would have avoided the death spiral if they had a variable costing system.

Cost-Volume-Profit Analysis: A study of the relationships between sales volume, expenses, revenue, and profit. Explicitly assumes that the incremental cost of any unit of activity is exactly the same as the incremental cost of any other unit.

Death spiral: A self-perpetuating processes in which a firm stops producing items with negative full-cost margin even though they have a positive contribution margin; this reduces the margin of remaining
products leading to further product cancellations and profit decreases. Exacerbated by cost-based pricing; mitigated by capacity costing.

**Avoidable fixed costs**: Costs that will be eliminated if a product line is eliminated. These costs are "fixed" in the sense that marginal increases in activity do not increase these costs, but they are "avoidable" in the sense that reducing production to 0 will eliminate the cost.

**Unavoidable fixed cost**: Costs that will remain--and be borne by remaining product lines--even if a product line is eliminated.

**Cost-based pricing**: Basing selling price off of reported costs. Can exacerbate the death spiral.

**Capacity costing**: Costing system that applies costs of used capacity to products, and treats the costs of unused capacity as period costs. Useful in avoiding the death spiral.

Theoretical capacity: The amount of total driver use under optimal (and unobtainable) conditions. The maximum driver use possible.

**Practical capacity**: The amount of total driver use assuming competent management, but recognizing that capacity is limited by a number of unavoidable factors, such as machine downtime and variation in product yields.

**Cost of unused capacity**: The amount of MOH that is underapplied simply due to the fact that the firm was not operating at theoretical (or practical) capacity. Firms that use capacity costing report this as one component of the adjustment to CGS.

**Idle capacity**: The portion of capacity that is unused (idle) for any reason.

**Surge capacity**: The component of capacity that is left idle because of variations in customer demand. Ice cream factories maintain surge capacity to handle high summer demand; that capacity is idle during the winter.

**Buffer capacity**: The component of capacity that is left idle because of variations in output. High-tech firms supplying just-in-time customers maintain buffer capacity for processes that have unpredictable yields, in order to ensure that they can produce enough successful units to meet customer demand each day.

**Strategic capacity**: The component of capacity that is left idle because some manager has made a strategic decision to invest in capacity that cannot be put to immediate use. Managers often invest in capacity for markets where they think that demand will rise in the future.

**Contractual capacity**: The component of capacity that the firm has agreed to maintain idle in case the customer wants it quickly put to use. The government may pay a car manufacturer to have a larger plant than necessary in order to be prepared in the event of war.
Glossary

**Excess capacity**: Capacity that is idle because no one (neither customers nor managers) is demanding it. Charging the cost of excess capacity to customers may lead to the death spiral.

**First Commandment of Capacity Costing**: Thou shalt not charge the costs of unused capacity to those who demand it.

**Second Commandment of Capacity Costing**: Thou shalt not charge the costs of excess capacity to customers.

**Nested Capacity**: Production capacity is determined by a hierarchy. Top level capacity allows capacity at a lower level which allows capacity at lower levels. Mowing square meters of a client’s lawn requires transportation to the client which requires equipment.

**Higher-level Activity**: In order to engage in lower-level activities, higher-level activites are required. Traveling to an airport is a higher-level activity than flying miles.

**Cost Hierarchy**: A depiction of the levels of costs and activities that reflects the nature of nested capacity in a business operation. The textbook cost hierarchy includes facility level costs (the top), product level costs, batch level costs and unit level costs (the lowest).

**Activity-Based Costing**: A method of allocating overhead that uses recognizes levels of costs (a cost hierarchy) that are variable with respect to types of activities other than unit production volume.

**Facility-Level Costs/Activities**: Costs and activities that enable the overall operations of a business. The level closest to the traditional notion of fixed costs.

**Product-Level Costs/Activities**: Costs and activities that enable the production of a type of product or service, such as adding a stock-keeping unit (SKU) or a new offering in a customer catalog. Variable with respect to the number of product/service types or SKUs, but fixed with respect to the number of production batches or units.

**Batch-Level Costs/Activities**: Costs and activities that enable the initiating of a production run or other batch-type activity. Variable with respect to the number of production runs or batches, but fixed with respect to the number of units produced in the run or batch.

**Unit-Level Costs/Activities**: Costs and activities associated with creating one additional unit of a product or service, once the production run has begun. The level closest to the traditional notion of variable costs.

**Activity Trigger**: An event that initiates a cluster of activities, commonly used as a driver for the pool of associated activity costs. Addition of a product line triggers a cluster of product-level activities; start of a production run triggers a cluster of batch-level activities.
**Volume-Based (Traditional) Costing:** A cost system that includes only unit-level drivers. Volume-based systems place too much cost on high-volume products and services, and too little cost on low-volume products and services, which tend to include high product- and batch-level costs spread over few units.

Volume-based cost driver: A cost driver that is closely associated with production volume, such as direct-labor hours or machine hours.

**Activity Chain:** A sequence of activities that form a hierarchy in which each level enables the subsequent activity. Firms may have multiple activity chains.

**Activity Dictionary:** A comprehensive list of the activities required for production.

**Bill of Activities:** A list of the activities required for the production of a particular product or service.

**Activity Clusters:** A set of activities that occur together in response to an activity trigger.

**Activity-Based Management:** The use of Activity-Based Costing information to reduce the cost of activities and the frequency of non-value-added activities.

**Activity-Based Pricing:** The use of Activity-Based Costing information to generate prices that more accurately reflect the cost of the value-added activities customers desire.

**Non-Value-Added Activity:** An activity whose value to the customer is less than its cost to the producer, and therefore reduces the supply chain surplus. Non-value-added activities may be (1) unnecessary and dispensable or (2) necessary but efficient and improvable.

**Supply Chain Surplus:** The value of a product or service to an end customer, minus the summed costs of all of the suppliers in the chain of companies needed to provide the product or service.

**Distributive Bargaining:** A bargaining setting in which the parties are restricted to deciding how to allocate a fixed economic surplus (dividing a fixed pie). Sharing information provides little economic benefit in purely distributive settings.

**Integrative Bargaining:** A bargaining setting in which some outcomes can make both parties better off. Effective integration creates economic surplus (makes the pie bigger). Sharing information (through activity-based pricing or even "opening up the books" can be beneficial in integrative bargaining settings.

**Credit and Blame:** A causal attribution about how individual actions caused good or bad observed outcomes. Very difficult to assign!

**Authority:** The right to make and implement decisions. People with authority have responsibility, but people with responsibility may not have authority.

**Responsibility:** The duty to know how well your portion of the organization is performing, explain deviations from expected performance, and propose responses.
Accountability: Having rewards or punishments associated with a performance measure. A manager is accountable for a line item if the status of that item will determine bonus, promotion, reprimand, or termination.

Responsibility Center: Semi-autonomous organizational units that are (1) directed by a single party and (2) evaluated based on its own set of managerial reports.

Isolating Responsibility: Identifying one party to be responsible for an individual line item on a report. The responsible party is said to "own" the line item.

Investment Center: A responsibility center evaluated on the basis of an income statement and a balance sheet. The center closest to a stand-alone business, evaluated on the basis of return on assets and well as profit.

Profit Center: A responsibility center evaluated on the basis of an income statement, but not a balance sheet.

Revenue Center: A responsibility center evaluated on the basis of only the revenue portion of an income statement.

Cost Center: A responsibility center evaluated on the basis of only the cost portion of an income statement.

Transfer Price: The price assigned to a transfer.

Internal Supply Chain: A supply chain consisting entirely of responsibility centers within an organization.

Related Parties: Two parties who serve the same principal, such as two responsibility centers within an organization.

Normal Cost Transfer Price: A transfer price equal to the normal absorption cost of a good or service transferred. Sometimes called "normal absorption cost" or "normal full cost" transfer price for clarity. Useful for tracking the flow of costs.

Standard Cost Transfer Price: A transfer price equal to what a good or service should have cost according to some set of performance requirements. Useful for isolating responsibility for production efficiencies.

Make-Or-Buy Decision: A decision of whether to produce a product or service inside the organization (often by another responsibility center) or purchase it from an outside supplier.

Variable Cost Transfer Price: A transfer price equal to the (normal or standard) variable cost of a good or service transferred. Useful for assisting make-or-buy decisions when the transferring center has excess capacity.

Market Transfer Price: A transfer price equal to the market price for a good or service transferred. Useful for assisting make-or-buy decisions when the transferring center has NO excess capacity.
**Cost-Plus Transfer Price**: A transfer price equal to the normal or standard absorption cost of a good or service transferred, plus a mark-up to allow the transferring center to report gross margin. Useful in isolating responsibility for value creation.

**Negotiated Transfer Price**: A transfer price determined by a negotiation between the parties involved in a transfer. Useful in isolating responsibility for value creation.

**Decentralization**: Allowing directors of responsibility centers the authority to make their own decisions.

**Centralization**: Retaining authority for decision-making at the highest levels of the organization.

**Specific Knowledge**: Knowledge that is costly to communicate to others. May include expert knowledge as well as knowledge of rapidly changing conditions. A good reason to decentralize decision-making.

**General Knowledge**: Knowledge that is easy to communicate to others. A good reason to centralize decision-making.

**Budget**: A financial plan that projects financial statement accounts, operational performance and expenditures.

**(Budget) Variance**: A deviation of actual performance from budgeted (planned) outcomes.

**Operational Projections**: Budget Projections describing the number of units to be produced and sold, and the resources required to achieve those levels.

**Financial Projections**: Financial statements and key performance indicators that would result from achieving operational projections.

**Cash Flow Projections**: A projection of cash inflows outflows and balances every day of the year. Make sure that required cash reserves are available every day.

**Performance Standards**: Levels of performance required to achieve operational projections.

**Periodic Budget**: A budget for a period that does not overlap with the previous or next budget. (A new 12-month budget every year)

**Rolling Budget**: A budget that is updated more frequently than the length of the budget. (A new 12-month budget every month).

**Budget Creep**: A tendency for each part of an organization to demand more resources every time the budget is revised.

**Zero-Base Budgeting**: A budget process requiring every part of the organization to justify their resource requests relative to a baseline of no resources at all, rather than a baseline of last budget's resources. A means of avoiding budget creep.
**Glossary**

**Pull Budgeting**: Budgets that place priority on demand for the company's output, rather than feasibility of supplying output. Places budgetary power primarily in the hands of sales and marketing.

**Push Budgeting**: Budgets that place priority on feasibility of supplying output over demand for output. Places budgetary power primarily in the hands of production and operations.

**Top-Down Budgeting**: Budgets that place priority on owners' demand for performance, rather than subordinates' ability to supply performance. Places budgetary power primarily in the hands of top corporate management.

**Bottom-Up Budgeting**: Budgets that place priority on subordinates' ability to supply performance, rather than owners' demand for performance. Places budgetary power primarily in the hands of lower management.

**Stretch (Performance) Standards**: Challenging levels of performance that are attainable, but more likely to be missed than met. Ideal for motivating workers and achieving high performance; not as good for planning or responding to surprises.

**Practical (Performance) Standards**: Levels of performance that are expected to occur, and are as likely to be missed as met. Useful for responding to surprises because variances are zero unless something surprising happens; not as good for motivation or planning.

**Minimum (Performance) Standards**: A level that is very likely to be achieved except through unacceptably low input. Useful for cash management (in which poor performance can lead to business failure) and for avoiding unpleasant surprises for investors. Inhibits management by exception because expected performance leads to favorable variances.

**Management by Exception**: Directing attention to large variances and deviations from budget projections. Assumes that conformity to the budget is unsurprising, and that only deviations need to be noticed, explained and responded to.

**Sandbagging**: Surrupetiously adjusting forecasts and estimates from true expectations in an attempt to generate budgetary slack. Sandbagging is an INPUT into the budget process; budgetary slack is an OUTPUT.

"**Budgetary slack**: Receiving more resources than needed to achieve a stated level of performance. OR Targets and standards that are easier to achieve than intended. Budgetary slack is an OUTCOME of the budget process."

**Paying People To Lie**: Michael Jensen's description of the fundamental tension between wanting accurate information from managers in the budget process and wanting to use that information to determine the managers' compensation (inducing them to lie). Sandbagging is the most common budget "lie" in an attempt to get budgetary slack.

**Pay Floor**: A minimum level of pay (base salary)
Glossary

Pay Ceiling/Cap: A maximum level of pay.

Hurdle/Hurdle Bonus: A hurdle is a level of performance resulting in a hurdle bonus; a hurdle bonus is the bonus received for passing a hurdle.

Kink: A point at which the slope of a pay-for-performance contract changes sharply. Hurdle bonuses, pay floors and pay ceilings all generate kinks.

Convex Compensation: A pay plan in which the incremental reward for improving performance increases as the level of performance increases. Induces risk-seeking and anti-smoothing.

Concave Compensation: A pay plan in which the incremental reward for improving performance declines as the level of performance increases. Induces risk-avoidance and smoothing.

Big Bath: Choosing to report even worse performance once performance is already poor, in order to maximize compensation or performance evaluation in future periods. Especially common when performance is just below a kink in the compensation plan. Can be done by accelerating expenses or delaying revenue.

Smoothing: Managing operations or reporting in order to generate similar performance across periods. Induced by concave compensation schemes.

Jensen's Dream Contract: A pay contract that is (a) entirely divorced from budgetary information and (b) perfectly linear, with no kinks or curves. Makes the budget process more honest, but (a) reduces controllability in pay, (b) increases workers' pay risk, © forces high compensating differential to attract risk-averse workers, and (d) makes it hard for firm to project compensation costs.

At-Will Employment Contract: A contract that (1) can be cancelled by either party on short notice and (2) imposes enough behavioral, financial or relationship control to be an employment contract.

Independent Contractor: “The general rule is that an individual is an independent contractor if the payer has the right to control or direct only the result of the work and not what will be done and how it will be done.” (IRS)

Deliverables: The outputs or outcomes to be delivered under an independent contractor agreement.

Behavioral Control: A business's instructions about when, where, and how to work. The IRS uses behavioral controls as evidence that a worker is an employee, not independent contractor. When and where to do the work. What tools or equipment to use. What workers to hire or to assist with the work. Where to purchase supplies and services. What work must be performed by a specified individual. What order or sequence to follow when performing the work.

Transactions Costs: The costs of writing and monitoring contracts. High transaction costs make long-term employment contracts more preferable than independent contractor arrangements.
Glossary

**Measurability:** The ease with which a firm can describe an objective and legally enforceable set of proxy measures to determine payments for deliverables.

**Standardizability:** The ability of the firm to reuse the terms of a contract for a deliverable.

**Granularity:** The ability of a deliverable to be provided independently of any other deliverables.

**Capital Intensity:** The dollar value an independent contractor would need to invest, per dollar value of labor.

**Specificity of Investment:** The ability of an independent contractor to repurpose their investment to be an effective contractor for another organization.

**Crowdsourcing:** Offering small amounts of money to many outside contractors for small amounts of standardized work from each. Examples include Wikipedia and Amazon Mechanical Turk

**Gamification:** The process of incorporating game elements into employee or consumer decisions. Examples: Avatars; 3-D settings; narrative content; reinforcement; reputations, ranks and levels; marketplaces and economies; competition under rules; teams; parallel communication systems; time pressure.

**Standard costing system:** A costing system that reports what it SHOULD have cost in order to produce a cost object, rather than what it ACTUALLY cost.

**Standard:** A target drawn from the budget that determines the dollars transferred into an account. Standards for costs determine the debit to inventory when goods are produced; standards for revenue determine the credit to (standard) revenue.

**Variance:** For costs, the difference between the amount applied to inventory according to the budget and the amount actually incurred. Analogous definition for sales (difference between recognized and actual revenue).

**Drill-down Analysis:** The process of splitting total deviation from budget into as many different (meaningful) variances as possible. Variances are typically split by cost behavior (fixed vs. variable), by resources consumed (labor, material), by purpose of expenditure (product line, customer) & by time and place (e.g. night shift at Altoona plant)

**Favorable:** A variance that adjusts CGS downward, and therefore adjusts income upward.

**Decomposition:** Splitting the total difference between actual and standard into two parts. For variable costs, the usual decomposition splits the difference into a variance driven by input price per unit of input and another driven by quantity of input used for the amount produced. For fixed costs, the decomposition splits the difference into a variance driven by total amount spent and a variance driven by the volume of production.
Glossary

Isolation of Variance: Creating variances that are specific enough as to be "owned" by a single responsible party.

**Unfavorable**: A variance that adjusts CGS upward, and therefore adjusts income downward.

**Direct Materials Purchase Price Variance**: Over- or under-application of materials cost caused by actual materials cost per amount purchased deviating from standard cost per amount purchased. \((AP-SP) \times AQ\)

Direct Materials Quantity Variance: Over- or under-application of materials cost caused by actual material use deviating from standard use allowed given amount produced. \((AQ-SQ) \times SP\)

**Direct Labor Efficiency Variance**: Over- or under-application of labor cost caused by actual labor hours deviating from standard hours allowed given amount produced. \((AQ-SQ) \times SP\)

**Direct Labor Rate Variance**: Over- or under-application of labor cost caused by actual wages per hour deviating from standard wages per hour. \((AP-SP) \times AQ\)

**Variable Overhead Spending Variance**: Over- or under-application of variable overhead cost caused by actual variable OH/driver unit deviating from the predetermined rate/driver unit \((AP-SP) \times AQ\)

**Variable Overhead Efficiency Variance**: Over- or under-application of variable overhead cost caused by actual driver use deviate from driver use allowed for the achieved level of production \((AQ-SQ) \times SP\)

**Fixed Overhead Budget Variance**: Over- or under-application of Fixed OH due to a deviation of total FOH from the budget Actual FOH - Budgeted FOH

**Fixed Overhead Volume Variance**: Over- or under-application of Fixed OH due to a deviation in production volume from the volume assumed in the static budget Budgeted FOH - Applied FOH

Consequentialist Ethics: The view that the morality of an action is determined by its consequences

**Rules of Engagement**: Rules that determine what forms of measure management are allowable or unallowable.

**Deontological (Duty-Based) Ethics**: The view that the morality of an action is determined by whether the action violates a duty.

**Occupational Fraud**: Use of one's occupation for personal enrichment through the deliberate misuse or misapplication of the employing organization's resources or assets.

**Asset Misappropriation**: Schemes in which the perpetrator steals or misuses an organization's resources. One of the three forms of occupational fraud. The most frequent form, but least expensive per incident.

**Corruption**: Employee's use of his or her influence in business transactions in a way that violates his or her duty to the employer for the purposes of benefitting him- or herself or someone else. One of the three forms of occupational fraud.
**Glossary**

**Fraudulent reporting**: Intentional misstatement or omission of material information in a performance report. One of the three forms of occupational fraud. The least frequent form, but most expensive per incident.

**Enterprise Corruption**: Aiding an enterprise in the pursuit of criminal goals. Enterprise corruption is NOT occupational fraud, because the employer is not the victim.

**The Fraud Triangle**: The three elements that lead to occupational fraud. Pressure, Opportunity, Rationalization

**Pressure**: Part of the fraud triangle. What problems will fraud solve if I don't get caught? Financial, personal or performance downturns are more likely to lead to fraud than simply a desire for improvement.

**Opportunity**: Part of the fraud triangle. Can I commit the fraud and get away with it? Opportunity generally arises from weak internal controls

Impact on Audience: The direct benefit or harm inflicted upon the people relying on a report due to their being deceived. More harm implies the lie is more unethical.

Skimming: A form of asset misappropriation in which cash receipts are taken before they are recorded.

**Rationalization**: Part of the fraud triangle. Can I live with myself if I commit fraud (even if I get away with it)? Reframing fraudulent behavior to make it seem acceptable. Everyone is doing it. The victim doesn't deserve it, or isn't harmed. I am doing this for noble purposes. Also called "moral disengagement"

**Lie**: An assertion, the content of which the speaker believes to be false, which is made with the intention to deceive the hearer with respect to that content. Assertion, intentional, false, deceptive.

**Principle of Veracity**: A strong moral presumption against lying. If you are considering lying, you must arrive at a decision that would be acceptable to a reasonable public.

**Impact on Practice of Veracity**: The effect of a lie on the likelihood that future speakers will lie and future audiences won't be able to trust similar reports by others. An externality; the greater the harm on the practice of veracity, the more unethical the lie.

**ManagementAssertions**: Representations by management that are implied by the way events are reported by their accounting system. Assertions may be about (1) classes of transactions and events, (2) account balances and (3) presentation and disclosure.

**Occurrence**: A management assertion that transactions and events that are recorded have occurred. Usually called "existence" for balance sheet item

**Completeness**: A management assertion that all transactions and events have been recorded. For balance sheet items, that all assets, liabilities and equity interests have been recorded.
Accuracy: A management assertion that all recorded amounts have been measured accurately. For balance sheet items this becomes a matter of valuation and allocation.

Classification: A management assertion that transactions were recorded in the right account.

Cutoff: A management assertion that transactions were recorded in the right period.

Separation of duties: Separating operational, custody, record-keeping and authorization duties to ensure that no one person can commit occupational fraud.

Behavioral Bias: Systematic error in judgment or decision-making—usually unintentional

Cold Bias: Biases arising from limited processing ability (small brains)

Hot Bias: Biases arising from a defense mechanism to protect a favorable self image of being moral and capable (strong emotions)

Heuristic: A rule of thumb that usually helps judgments and decision-making, but leads to predictable errors (biases) in certain situations

Availability Bias: Overestimating the likelihood of events and facts that are easier to retrieve from memory (more available). Arises from overuse of the availability heuristic: Events and facts that are easier to retrieve from memory are more likely to be true.

Halo Effect: A form of overgeneralization: Your impression about one feature affects your impressions about other features. Across traits within a person, or across time. "If she's smart, she's also good with people." "Give a man a reputation as an early rise and that man can sleep until noon."

Planning Fallacy: People predictably underestimate time to complete tasks. Avoid by taking an outside view.

Motivated Reasoning: People analyze data in ways that justify their wishes--accepting good news uncritically, ignoring bad news or reading it carefully enough to find ways to discount it. A cause of wishful thinking

Fundamental Attribution Error: A tendency to attribute other people's actions to what kind of person they are, rather than to situational variables. In most cases, circumstances trump personalities. People generally behave similarly in similar circumstances.
I hope you’ve learned to see the world through new eyes, to distinguish what counts from what is counted, and to help your organization devise a better “nose” with which to sniff out margins and efficiencies, no matter how much intra-organizational conflict you might face. Enjoy your free lunches as you find them!

Rob