

# Designing Incentive Plans: New Insights from Academic Research

---

Incentive compensation has been one of the most actively studied topics in economics, accounting and management research in the past two decades. An active and fruitful interplay between theory and empirical research has generated a large, mature and robust literature. Much of that literature is technical and difficult for the practitioner to access, but the insights are practical. Designing an effective incentive plan will always require judgment and understanding of individual circumstances. This article describes how modern research can provide rigor and structure for the design challenge.



**Michael Gibbs, Ph.D.**  
University of Chicago Booth School of  
Business

## TWO CONDITIONS FOR AN EFFECTIVE INCENTIVE PLAN

It is helpful to break design of an incentive plan into two components and think about each separately, in a specific sequence. This approach is almost always an effective way to organize thinking about incentives. The two are:

- 1 | Effectiveness of the evaluation.** Does the evaluation reflect the employee's efforts, actions and decisions with reasonable strength and accuracy?

**2 | Relationship between the evaluation and rewards.** Is the evaluation tied reasonably strongly to rewards that the employee values?

These may seem obvious, but it is surprising how many “incentive” plans violate one or both conditions. For example, many large firms use profit-sharing plans, but such plans violate both conditions for all but key high-level employees. The performance measure is divisional or firm-wide profit. This does not reflect the employee’s efforts well because it depends not only on the employee, but also on the rest of the organization, and a large number of other factors that are outside the employee’s control. An individual employee is unlikely to have any practical ability to affect the performance measure, unless he/she is at a very high level in the company. The signal-to-noise ratio is nearly zero, so such an approach is tantamount to using a lottery ticket as the performance measure.

Now consider the second condition. Even if the employee creates a large amount of value, the share that worker receives from his/her contributions is small. Suppose that employee works for a company with 1,000 employees, and participates in a plan in which 25 percent of profits are shared with employees as a year-end bonus. Suppose further that the employee has a good day, creating \$1 million in profits for your firm. What is his/her share? It is  $25\% \cdot \$1M / 1,000 = \$250$ . The employee’s “commission” is only 0.00025. The bonus will be larger than \$250 because the employee will also be rewarded for the contribution of the rest of the organization. However, that will be true no matter whether the employee works hard or not, so it is unlikely to provide incentive. A firm would never use such a trivial commission rate for salespeople, so it is surprising that similar logic is so seldom applied to other employees.

Thus, incentive plan design should start with the basics. Effort should be directed at making sure the evaluation reflects the employee’s effect on firm objectives, and that rewards vary with how the employee’s contributions vary — not other things.

### **Focus First, and Most, on Performance Evaluation**

Focus first on performance evaluation without thinking about how that will be tied to rewards. That should be considered second. There are several reasons to begin with a thorough analysis of evaluation.

First, evaluation is the most complex issue. Most of this article focuses on the subject for just that reason. How the evaluation is tied to actual rewards is a much simpler question. Second, the extent to which rewards will be tied to the evaluation depends on the properties and effectiveness of the evaluation. Understanding these properties leads to better judgment about how to use that evaluation for rewards. Third, performance evaluation is the most important of the components. Most problems in incentive plans derive from the evaluation. Moreover, evaluation affects motivation and performance even if it is not tied to rewards. An effective evaluation can improve intrinsic motivation by providing feedback, so that the employee can learn and continuously improve. Evaluation is an important way to

train and coach an employee, and build trust. Performance management matters to performance in several ways beyond incentives. Therefore the ideas about evaluation discussed here are useful even if there is no intention to provide incentives.

### Principles of Performance Evaluation

There are two keys to an ideal performance evaluation. First, good evaluation incorporates what the employee can control and filters out what the employee cannot control, to the extent that is practical. It incorporates all of the employee's contributions because its purpose is to reward, and therefore motivate, the employee to contribute in all possible ways consistent with the employee's job. That includes the direct effects of an employee's efforts and decisions, and also the indirect effects (e.g., through cooperation and coordination with colleagues). The ideal evaluation incorporates nothing else, because anything else is outside of the employee's control, and therefore imposes risk on the employee. People are risk averse. If the incentive plan rewards and punishes the employee in part for things beyond his/her control, this is costly for the employee. That cost, in turn, implies a cost to the employer. Research finds that incentives based on more noisy performance measures generally require firms to pay higher base salaries as a risk premium to employees.

The second consideration in evaluation is that an ideal evaluation motivates the employee to use information and knowledge on behalf of the firm. Employees usually have knowledge, which the supervisor does not, that is valuable to performing the job. The employee interacts with the customer, observes the machine in operation or knows the content and logic of a spreadsheet. Because the employee knows things the supervisor does not, there is a limit to how much direct guidance the supervisor can give the employee about how to perform the job. How can the employee be guided beyond that limit? By giving the right goals, and discretion and motivation to figure out how to best achieve those goals. This is especially important in more complex, uncertain and dynamic environments, where the situation is constantly evolving and performing the job gives the employee valuable information. One key to a good incentive plan is to tap into that knowledge via appropriate evaluation.

### Numeric or Subjective Evaluation?

There are two general approaches to evaluation: quantification with a numeric performance measure or subjective judgment such as a performance appraisal. All else equal, a numeric measure is preferred. A numeric measure allows the employee to know where performance stands at any given point in time. There is little room for argument, as such measures are usually outside of the discretion of the supervisor. By contrast, subjective evaluations are subject to a host of problems, such as favoritism and bias. They create legal risks for the firm. A good subjective evaluation requires supervisor time and effort. Finally, the employee

must have reasonable trust that the supervisor will be fair for a subjective evaluation to be effective.

For this reason, in designing an incentive plan it makes sense to begin by choosing the best available numeric performance measure and analyzing its strengths and weaknesses. To the extent that it is effective, a numeric performance measure can be given strong weight for rewards. To the extent that it has important weaknesses, subjective evaluation can be used to address those weaknesses.

## **PROPERTIES OF NUMERIC PERFORMANCE MEASURES**

No performance measure is perfect. The question is, “How serious are the flaws in the measure?” The greater the flaws, the less weight should the measure be given for incentives, and vice versa. Academic research on evaluation suggests four key properties of numeric performance measures, based on the performance evaluation principles described above. Those properties are described here. Any numeric measure can be evaluated along these dimensions.

### **Distortion**

The first key performance measure property is Distortion: the extent to which the effect of the employee’s actions on the performance measure does not correspond to the effect on firm value. Anyone who has worked with incentive compensation is familiar with the principle, “You get what you pay for.” That is, incentives motivate the employee to do those things that are included in the evaluation, and not those things that, though valuable to the firm, are not. For example, revenue ignores costs, so a revenue-based incentive does not motivate cost containment. Profit includes revenue and costs, but only for a single period, thereby distorting incentives away from caring about future profits. A manager with a bonus based on this year’s profits may defer maintenance, research and development, or other activities that require a cost today with the hope of cost avoidance or revenues in the future.

All numeric performance measures, with the possible exception of stock price, have some distortion: they give too much weight to some dimensions of the employee’s job and too little (or even zero) weight to others. The reason is that some dimensions of an employee’s job are easy to quantify well, while others are much harder or nearly impossible to quantify. Quality, service, innovation and many other intangibles fall into the latter category. Modern methods of performance evaluation, such as the Balanced Scorecard or Economic Value Added, have improved the ability of firms to quantify intangibles, but those methods are still imperfect. So, there is always some distortion in a numeric performance measure.

A general concern with incentive plans is that they may cause employees to “game” the system, acting in ways that increase compensation but do not improve firm value. Distortion is the first of three common types of gaming. (The other two, manipulation and timing of performance in the presence of all-or-nothing rewards, are described later.) This type of gaming is “baked into” the evaluation

because the distorted motivation from an imbalanced performance measure is predictable: the employee can be expected to focus too much on things that the measure incorporates and too little on things to which it gives less emphasis.

Many, perhaps most, common criticisms of pay for performance are caused by distorted performance measures. Kohn's (1975) classic "On the Folly of Rewarding A, While Hoping for B" illustrates the idea clearly in its title. Other criticisms, such as that pay for performance undermines intrinsic motivation or creativity, are also caused by distortion. Numeric measures focus on what is most easily quantified. The desire to avoid "uncontrollable risk" reinforces this because quantifying an employee's qualitative contributions is very difficult to do with accuracy. Therefore, numeric measures almost inevitably distort incentives away from dimensions of the job that are hard to quantify, even if they are important to firm value. Behavior that we associate with creativity and intrinsic motivation is, of course, usually difficult to quantify. Therefore, incentive plans often reduce creativity. That does not mean that such plans are not useful, as they create other benefits.

However, it is possible to design incentive plans so that they are less likely to undermine creativity and intrinsic motivation. First, numeric measures could be relatively broad and focused more on outputs (end results) than inputs (how to get those results). Second, subjective evaluation is often a much better way to guide the employee's motivation toward behaviors that are difficult to quantify. Third, consider including a floor on rewards to avoid punishing mistakes because that would make employees more risk averse.

Since essentially all measures are distorted, the question is how much a specific measure is distorted. The more distorted a performance measure, the less valuable it is for incentives, and vice versa.

### **Uncontrollable Risk**

Risk in an employee's work environment comes in two forms, uncontrollable risk and controllable risk. These are the second and third key properties of a numeric performance measure. In thinking about performance evaluation, it is important to distinguish between the extent to which the uncertainty can be controlled by the employee or is uncontrollable, because the two types of risk have opposite implications for incentive plan design. The distinction between controllable and uncontrollable risk is a relatively new and important idea in academic research on incentives (Prendergast 2002).

Uncontrollable risk corresponds most closely to the classic idea of risk – events that are random and the employee can do nothing about. (However, as is described in the next subsection, the issue is more subtle than random or not random.) It is often referred to as measurement error. To the extent that the environment is unpredictable in this way, the performance measure will reward or punish the employee for things that he/she cannot control. This is an important flaw in performance measures, though one that many firms do not seem to pay enough

attention to in designing incentive plans. Returning to the earlier example, the performance measures in profit-sharing plans are largely affected by factors outside the employee's control. The problem was highlighted by comparing it to lottery tickets. Even worse is stock price. Stock price may not distort incentives, which is a good feature for a performance measure. However, there are many factors that affect stock price that are beyond the control of even the CEO, let alone rank-and-file employees. It has a very high degree of uncontrollable risk, and for that reason is a very poor performance measure for most employees.

All performance measures have some uncontrollable risk in them – it is impossible to perfectly quantify an employee's many subtle contributions. As with distortion, the question is how much is acceptable. In analyzing a performance measure, evaluate how much risk the measure has: how much variance is there over a given period of time? Next, analyze the extent to which that risk is uncontrollable or controllable by the employee.

Note that the first two performance measure properties, distortion and uncontrollable risk, derive from the first principle, that an ideal evaluation measures what the employee can control and filters out what he/she cannot. To the extent that it does not incorporate all of the employee's contributions, the measure is distorted. To the extent that it includes the effects of other factors, the measure has uncontrollable risk.

### Controllable Risk

Risks, or their effects on firm value, are often partially controllable by the employee. To illustrate, consider a factory roof that collapses due to accumulation of snow after a heavy snowstorm. This event is random – the plant manager cannot control the weather. However, there are many reasons why it might be proper to hold the manager accountable for the collapsing roof anyway. There may have been precautionary actions that should have been taken to reduce the likelihood of a roof collapse, including inspection and maintenance. The manager may not have taken appropriate actions, perhaps because of a distorted performance measure that reduced incentives for expenditures in the short run. In addition, a snowstorm is a slow-moving event, so the manager may have been able to react to it as it happened (e.g., shoveling snow off of the roof before too much accumulated). Finally, it is likely that there were actions the manager could take after the roof collapsed to mitigate the damage (ensure safety, commence repairs, second-source production temporarily). In short, though the manager cannot control the weather, there are many ways in which he/she can control how the snowstorm affects firm value. A risk is controllable to the extent that the employee can control the effect of the risk on firm value, even if the risk itself is random.

Clearly, an ideal incentive plan will motivate the plant manager to foresee risks, take precautions, react as unexpected events unfold and manage their aftermath.

This illustrates why it is important to think carefully about whether environmental risk reflected in a performance measure is uncontrollable or controllable. Risks that are uncontrollable should be filtered out of the evaluation if possible. If that is not possible, the evaluation is less valuable and should be given less weight for incentives. By contrast, risks that are controllable should be included in the performance measure in order to motivate the employee to effectively manage those risks.

Controllable risk and the next performance measure property — manipulability — derive from the second principle, that an ideal evaluation motivates the employee to use his/her knowledge on behalf of the firm. As employees do their jobs, they observe real-time conditions that affect their work (e.g., how angry a customer is and the nature of the complaint). They have a deep and complex understanding of their work that their supervisors often do not have. An ideal performance evaluation holds the employee responsible for using information and knowledge appropriately. To the extent that a measure captures the employee's use of knowledge to improve firm value, the measure is more effective and should be given more weight for incentives. For example, output-based measures are more likely to capture the employee's use of information and knowledge more effectively than measures that reward an employee for performing activities in a specific way. The latter approach requires knowledge of the best method to achieve results. The former approach gives the employee a goal but does not specify how to achieve that goal. It therefore encourages the employee to use insights and creativity to figure out the best approach.

### Manipulability

The flip side of controllable risk is manipulability, the fourth key property of a numeric performance measure. In some cases the employee can use an informational advantage to improve a performance measure in ways that do not improve, or may even harm, firm value. Consider a health-care organization that measures surgeons by mortality rates (Prendergast 2002). The surgeon has an incentive to use his/her expertise in examination of patients to operate only on patients with relatively simple medical problems. By manipulating performance measure in this way, the surgeon improves his/her evaluation, but avoids the type of patients that the hospital actually wants to serve. In evaluating the effectiveness of a performance measure, one should consider the extent to which risks are controllable. One should then ask whether those controllable risks help the employee improve firm value or are more likely to be used to manipulate the measure. If it is the latter, the evaluation is less useful for incentives.

Manipulation is the second common type of gaming of an incentive system; distortion was the first. They are similar in that they involve the employee doing too much of some things, and/or too little of other things, to maximize the performance measure and incentive. However, it is useful to distinguish between them because these two problems are addressed differently, as will be discussed later.

## OTHER REMARKS ABOUT PERFORMANCE MEASURES

### Degrading of Performance Measures

An interesting implication of these ideas is that the quality of numeric performance measures can degrade over time (Courty and Marschke 2008). Suppose that a firm finds that a performance measure historically had high correlation with desired outcomes, and decides to give a bonus based on the measure. Two problems may arise. First, the new incentive may now distort the employee's behavior, so that the performance measure becomes a less useful signal of what the firm intended to measure. Second, the employee may manipulate the measure, making the situation worse. An example of both is the unfortunate consequences that sometimes result when a university tries to evaluate research productivity of professors. The goal is to motivate a good quantity of high-quality research. One potentially useful metric is the number of papers that a professor publishes. However, this metric ignores quality, and so it likely to distort the behavior of professors, who may begin to write a large number of low-quality articles. Worse, professors might try to manipulate the system – for example, setting up journals to publish their own research or that of their colleagues. The metric “number of published articles” seemed informative at first, but its quality declines as behavior changes and professors continuously improve their ability to game the incentive system.

### Performance Measure Scope

Four performance measure properties have been discussed: distortion, controllable and uncontrollable risk, and manipulation. One useful way to think about these properties in practice is a measure's scope: the extent to which it is narrowly focused or broadly defined. For example, profit is broader than revenue or cost because it incorporates both of those measures. A profit measure's scope could be expanded further along the dimension of time, say, three years instead of one. Economic Value Added (EVA) broadens profit by using the discounted present value of profits and incorporating the opportunity cost of capital into costs.

Scope illustrates that a tradeoff may exist between uncontrollable risk and distortion in performance measurement. Consider stock price, a measure with very broad scope (it is the market's estimate of the discounted present value of cash flow to shareholders). As discussed earlier, this measure includes the effects of everything the employee can do that affects firm value, and thus has little distortion. However, it includes a large amount of uncontrollable risk.

Because of this high risk, stock price is a problematic performance measure for most employees and is generally not used as such. Stock or option plans use stock price as the performance measure. Such plans seem a poor incentive for employees other than the most senior ones in the organization, because they violate both conditions for an effective incentive plan.

Even for CEOs, it is common to base some incentive on narrower performance measures, particularly earnings. The benefit of narrower measures is that they filter out a great deal of uncontrollable risk. Accounting earnings are far more controllable than stock price. However, earnings are also more distorted than stock price, since they are a short-term measure.

Similar tradeoffs may arise between controllable risk and manipulation when varying the scope of a measure. Random events are often partly controllable and partly uncontrollable. When this is the case, broader measures will tend to include more of both types of risk, while narrower measures will include less. Narrower measures also tend to be more manipulable. When a measure is affected by only a few factors, the effect of each factor is larger, and manipulating one has greater impact.

Putting these arguments together, the concept of scope may be useful in thinking through the tradeoffs between different performance measure properties. Once the aspects of performance that the evaluation is intended to motivate are chosen, there may be latitude in how narrowly or broadly to implement the measure, along several dimensions. Broader measures tend to have less distortion, more controllable risk, and be less manipulable, all of which make the evaluation more effective. However, they also tend to have more uncontrollable risk. Most measures narrow focus somewhat for this reason, creating some distortion, controllable risk and manipulation.

## **SUBJECTIVE PERFORMANCE EVALUATION**

It is very difficult to think of any job in which subjective evaluation or supervisor discretion does not play some role in incentives. Managers implicitly reward good performance through assignments, corner offices, promotions, etc. Formal plans must be adjusted from time to time, which involves discretion. Finally, many plans make explicit use of discretion through subjective evaluations or awarding discretionary raises and bonuses.

The problems with subjective evaluation are well known (Murphy 1992; Prendergast and Topel 1993; Murphy and Cleveland 1995; Prendergast 1999). Supervisor preferences and incentives may play a role in the evaluation, resulting in favoritism or manipulation of the system to reduce compensation costs. Subjectivity may expose the firm to lawsuits alleging discrimination or wrongful termination. Subjective evaluation can also distort incentives or manipulation. The employee might behave in ways the supervisor prefers rather than ways that improve firm value. The employee might complain and make it difficult for the supervisor to give poor evaluations, resulting in evaluations that are too lenient or too concentrated (leniency and centrality biases). Supervisors might also have hindsight bias, holding the employee responsible for factors known by the supervisor *ex post*, but which the employee did not know when he/she performed the work. Finally, subjective evaluations impose risk on employees to the extent that they do not trust their supervisor to be fair.

Despite these problems, subjective evaluations must have benefits because they are so ubiquitous. A small literature review considers these benefits (e.g., Gibbs et al. 2004). The basic conclusion is that subjective evaluations should be used to address limitations of the available numeric performance measures. For this reason, a good incentive plan design will usually include some subjective evaluation (possibly in a second bonus), and use it consciously to address imperfections in numeric measures. It can be used to help the overall evaluation (numeric and subjective combined) come closer to meeting the two principles of performance evaluation described earlier. Subjectivity can improve the extent to which the evaluation captures the effects of all actions the employee can take and help filter out the effects of factors that the employee cannot control. Subjective evaluation can also help motivate the employee to use information and knowledge to take initiative, be creative and effectively react to a dynamic work environment on behalf of the employer.

A supervisor can reduce distortions by incorporating other dimensions of performance, such as quality and customer satisfaction, into the overall evaluation. The supervisor can give such dimensions of the job relatively more weight in the overall evaluation, balancing incentives for the employee to focus on different aspects of the job. For example, the plant manager might be given a formal bonus based on profits, but a second bonus or promotion opportunities might be based on a subjective evaluation. The profit-based bonus will distort incentives, possibly toward deferring maintenance, but the supervisor can use the subjective reward to include consideration of factors that are difficult to quantify.

Subjective evaluation often allows for subtle treatment of controllable versus uncontrollable risk. In the example of the collapsed roof, it is possible that the plant manager failed to take reasonable precautions, including deferring maintenance, and disaster ensued. However, it is also possible that the manager did all of the right things in an attempt to manage risks correctly, but the roof still collapsed. In the former case, the manager should be punished or even fired, but in the latter case the manager might even be rewarded. A numeric performance measure will fail to distinguish between these cases. An ex post subjective evaluation, however, may be able to do so. The supervisor can assess the manager's actions taken prior to, during and after the storm, and decide whether they were appropriate. The supervisor can take into account the information that the manager had at the time decisions were made, which is likely to be less than is now known. What the manager knew at the time is more likely to be controllable, while what was learned subsequently was uncontrollable. A wealth of factors may be relevant for fully assessing the manager's responsibility. It is very unlikely that a numeric measure chosen ex ante will incorporate all such considerations and adjust to the unique circumstances of a given situation. Subjective evaluations seem almost inevitable in order to provide an effective overall evaluation.

Discretion also provides a way to reduce manipulation, because it allows for ex post settling up. If a performance measure indicates very good performance,

but the supervisor has information or observations suggesting that this was partly due to manipulation, he/she can reduce the evaluation accordingly. This threat can deter manipulation.

Discretion in incentive systems has additional benefits beyond mitigating weaknesses in numeric measures. One such benefit is flexibility. Changing a measure mid-year is possible, but at some cost. Subjective evaluations and discretionary rewards are easier to change because they are not codified formally. This provides adaptability of the incentive system to changing conditions or the fixing of errors in design.

Subjective evaluation can be an important source of training and continuous improvement. Consider again the plant manager with the collapsed roof. During the evaluation the supervisor can identify ways to improve job design, resources, information or training. The supervisor can coach the manager on what to learn from the experience, and how to do the job differently in the future.

Only a small empirical literature exists on the use of subjective evaluation or discretion in incentive systems (e.g., Ittner, Larcker, and Meyer 2003; Gibbs et al. 2004, 2009), presumably because quantifying the concepts is difficult. The evidence is consistent with the uses of discretion described earlier. For example, subjective evaluations are used to reduce short-term incentives from numeric measures and to motivate cooperation across organizational units. Implicit punishments, such as threat of termination, are used to punish detected manipulation. Implicit rewards are used to award a manager who took appropriate actions but whose numeric performance is too low due to uncontrollable factors.

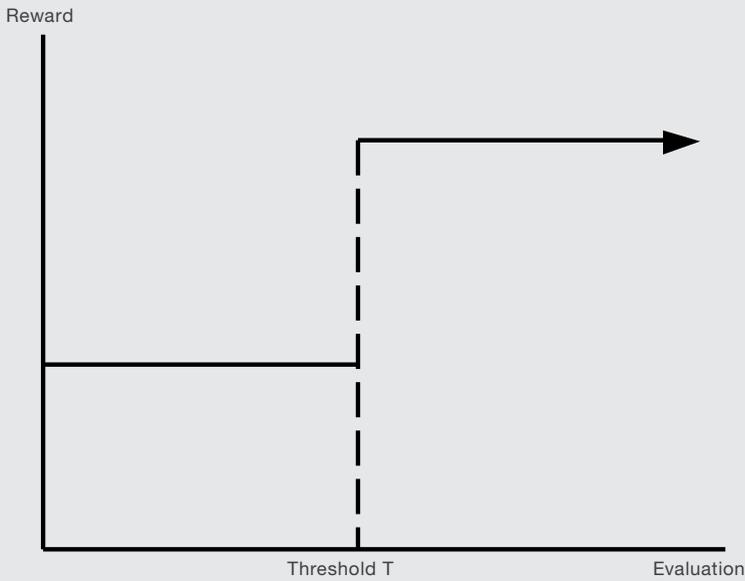
## **RELATIONSHIP BETWEEN THE EVALUATION AND REWARDS**

The second component of an incentive plan is how strongly rewards vary with the evaluation. Figures 1a and 1b illustrate two common approaches. In each, the horizontal axis represents the evaluation, which could be numeric or subjective. If it is subjective, the incentive plan might quantify it, such as a rating on a scale of 1-5 or A-E. If the evaluation is subjective but no formal score is applied, think of the horizontal axis as graphically representing the supervisor's judgment of the employee's performance. The vertical axis is the reward. If the incentive involves a formal bonus, this is simply the payout. However, many rewards are implicit — job assignments, access to training opportunities, promotions, etc. In such cases the vertical axis graphically represents the value of the implicit reward to the employee.

The important point is that a better evaluation must lead to greater expected rewards — that is the second condition of an effective incentive plan described at the beginning of this article. Therefore, for this component to motivate, the general shape of the reward-evaluation relationship should be a positive slope of some kind — higher rewards at higher performance. Beyond that basic point, many “shapes” of this relationship are possible. Figure 1a corresponds to a case where the reward is all-or-nothing. For example, a firm might offer a lump-sum

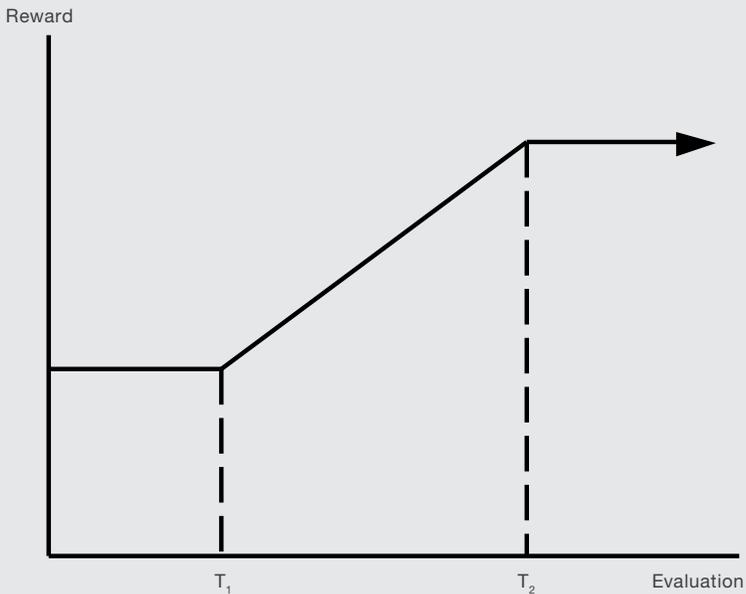
**FIGURE 1A** Reward is All or Nothing

Lump-sum bonus for performance above threshold T. Promotion-based incentives have the same shape.



**FIGURE 1B** Reward Varies with the Level of Performance

Base compensation is paid for performance below  $T_1$ ; a variable reward is paid for performance above  $T_1$ , up to  $T_2$ .



bonus to a salesperson who meets or exceeds a target. Promotions have the same type of shape, as they are awarded to employees who exceed some criteria, but not to those who do not.

Figure 1b corresponds to a reward that varies with the level of performance. Between the two threshold levels of performance  $T_1$  and  $T_2$ , the reward increases

as the evaluation improves. An example is a bonus that is calculated as a constant multiple of a performance measure for performance levels between the thresholds. Another example might be a discretionary bonus in which the supervisor pays larger bonuses for better subjective evaluations, again between the thresholds. The shape illustrated in Figure 1b also has a floor – a reward is given only for performance above the minimum threshold T1 – and a cap – the reward has a maximum at performance equal to T2. Reward schemes may have either a floor or a cap, or neither.

Most bonus schemes look something like one of these two shapes, though simple linear pay-performance relationships are most common (Murphy 2000; Gibbs et al. 2009). In one study, about 2 percent of plans used lump-sum bonuses, 6 percent had a floor and 2 percent had a cap. Caps and lump-sum rewards are more common in cases where an employee is offered two or more bonuses. Caps also seem to be more common in executive incentive contracts. Let us briefly consider some implications of these general ways to tie pay to performance.

The key idea is that the steepness of the relationship between the reward and the evaluation is what drives incentives. Rewards must vary when performance varies. If an employee is paid pure salary, the relationship is flat, and compensation does not improve motivation. The steeper the relationship, the stronger the motivation.

### Lump-Sum Rewards

In the approach illustrated in Figure 1a, the reward-evaluation relationship has dramatically different slopes, depending on the level of performance. For low or high performance, the relationship is flat, just like simple salary. Variations in performance have no effect on rewards. For performance close to the threshold, the slope is vertical (infinite), so an employee whose performance is near the threshold is likely to have very high motivation.

Two common problems arise from this shape. First, if the firm sets the threshold too low or too high — the goal is too easy or too difficult — the employee is likely to have very little or no incentive. This was famously described as the problem of “goldbricking” and documented by Roy (1952) in one of the most famous articles ever written on incentives. It is a common concern with this type of reward shape, as it is often difficult to set a threshold so that it strikes the right level of difficulty.

The counterargument is that motivation is very high if the threshold is set appropriately. That is true, but then the second problem arises. Since the reward is all-or-nothing, the employee has a strong incentive to increase the performance evaluation by any means possible. That may generate stronger effort, but it also creates maximal incentives to manipulate the measure. This is the familiar problem of “sandbagging” or “channel stuffing” in a sales incentive contract, in order to either delay sales into the next year (if this year’s target has already been achieved), or accelerate sales into this year (if the target is nearly achieved and the end of the year looms). Indeed, academic evidence finds that dramatic changes in the slope of

the reward-evaluation relationship increase the likelihood of manipulation (Healy 1985). This is the third common form of “gaming” of an incentive system. While Figure 1a shows the most extreme version, any time that the reward-evaluation relationship changes abruptly (including at the two thresholds in Figure 1b), there is potential concern about manipulation. This does not mean that such shapes should not be used. However, it is a drawback of such shapes. When they are used and the employee’s performance is in the range where the shape changes significantly, the firm should be more diligent than usual in trying to detect manipulation. In addition, the general lesson is that smoother reward-evaluation relationships, such as the one in Figure 1b, tend to be preferred because they are much less subject to manipulation.

Given this analysis, why might a firm use a shape like that in Figure 1a? For bonus schemes, they should usually be avoided. One important exception is when it is desirable that an employee’s performance falls within a narrow range – the firm wants to motivate the employee to hit a target, but not to exceed it by too much. An example might be a plant manager, when average costs of production rise significantly for output levels that are below or above a certain level that corresponds to the plant’s most efficient rate of output. However, such circumstances are rare.

As an important side note, if the firm does not desire the employee’s performance to be too far above some level of the evaluation, then it must be the case that the evaluation imperfectly reflects the employee’s actual value to the firm. Otherwise, the firm would just want to encourage even greater “performance.” Put another way, because performance evaluation is imperfect, performance evaluations do not always measure performance, particularly for extreme values. Numeric measures and subjective evaluations are proxies for the employee’s actual contributions. For most ranges of performance, they will correlate well with contributions. When they are very low or very high, however, they are more likely to reflect good or bad luck, measurement error, flaws in the method of measurement or judgment, or manipulation of the metric by the employee.

The second reason why a firm might use a reward-evaluation shape similar to Figure 1a is that it might not have a choice. Promotions have this form: generally they imply a discrete increase in compensation, benefits, perks and other job characteristics. Indeed, academic evidence suggests that over the long run for clerical and white-collar workers, promotions are the most important source of increased compensation other than switching employers. By implication, since promotions are usually based on performance, they are one of the most important forms of incentive compensation for many employees. When an employee is promoted, he/she has a new line on the resume, and the firm has publicly stated that it believes the employee to be more valuable than before the promotion. In most cases, such an employee now has a greater value on the labor market. Therefore, it is not surprising that firms give lump-sum increases in rewards upon promotion; they often have to

in order to remain competitive in retaining the promoted employee. There seems little that a firm can do to avoid such situations. However, there is an important implication. Employees who have a reasonable chance of earning a valuable promotion may have good incentives built into the system, even if they receive no formal incentive such as a bonus. By contrast, employees who have little hope of further promotion (e.g., those who have topped out at the highest hierarchical level they are likely to achieve) will not have promotion-based incentives. The issues faced in such cases are well known, and are often referred to as the problems of “dead wood” or the “Peter Principle.” An interesting question is why so many compensation systems seem to ignore this problem; they do not substitute some other form of pay for performance for employees who are no longer in contention for promotion.

### Continuous Rewards and Performance Evaluation Properties

Now let us turn to Figure 1b. For the moment, ignore the floor and cap (assume that they are not there, so that the reward-evaluation relationship is a simple line). Recall that the slope of this line measures how rewards vary with evaluated performance. For example in a sales bonus plan, the slope is the commission rate. The goal is to offer a steeper relationship if possible. What factors affect this relationship and thus the ability to offer strong or weak incentives? The properties of performance evaluation (numeric, subjective or some combination) described earlier are the key. They describe the first condition for an effective incentive plan, effectiveness of evaluation.

For example, if the evaluation is less distorted, then motivating the employee more strongly does not lead to behavior that is too imbalanced, and the kinds of problems described in Kohn’s classic work are less of a worry. Similarly, if the evaluation includes a large amount of uncontrollable risk, putting strong weight (steep slope) on it for incentives imposes more risk on the employee because the employee’s rewards vary more with things outside of his/her control. Therefore measures with more uncontrollable risk should have a less-steep tie to rewards.

Controllable risk is a bit more complex. To the extent that the employee can foresee, react to, or adjust ex post to the effects of risks, the risk is controllable. In most cases, that is a good thing since the employee can use the information gained in doing the job to promote firm value. However, in some cases the employee can exploit the informational advantage to manipulate the performance measure, increasing compensation with no (or negative) contribution to firm value. Therefore, if the employee’s job entails a large degree of controllable risk that can be used to help the firm’s goals, the reward-evaluation relationship should be relatively steep to strongly motivate the employee. If the performance measure is relatively manipulable, the opposite is true.

This discussion illustrates why the first step in designing an incentive plan should be to focus on the evaluation, and why the properties of distortion, controllable

and uncontrollable risk, and manipulation are useful in thinking about which measures should be used, and how subjective evaluations should be conducted. Such analysis is a necessary input into how to tie rewards to the evaluation, so the reward-evaluation step logically comes second and is a separate component of an incentive system.

### Floors and Caps

An advantage of the shape in Figure 1b compared to Figure 1a is that there is continuous motivation for employees over a wider range of performance between the two thresholds. Furthermore, there is much less concern about manipulation for performance in that range. That said, the slope does change significantly at both thresholds. If an employee's performance is near T1, that worker may be motivated to manipulate performance to move into the range where some bonus is earned. If the employee's performance is above T2, he/she may begin to "gold-brick" (be less motivated) or "sandbag" (save some output for next year) since there is no additional reward from increasing performance. Given these concerns, why might a floor or cap sometimes be useful?

A benefit of a floor is that it provides some insurance to a risk-averse employee that he/she will not be punished (through lower rewards) for low performance. That is particularly valuable in cases where very low performance tends to be caused by bad luck (uncontrollable risk that adversely affects performance) instead of low effort by the employee. Providing some insurance to the employee is valuable in two ways. First, the employer may be able to pay a lower base salary since the employee will require less of a risk premium if the pay plan is less risky. Second, a floor has the effect of making an employee somewhat less risk-averse about the incentive plan. That, in turn, may allow the firm to use a steeper reward-evaluation slope, intensifying incentives for performance above the floor. More generally, in jobs where the firm wants to encourage risk taking and creativity (e.g., research and development, advertising), or wants to avoid punishing the employee for mistakes, a floor can be useful because it makes the employee less averse to taking some risks in performing the job.

Caps have a simple tradeoff. A cost of a cap is that the firm runs the risk of losing its star employees because they may hit their maximum reward before achieving their maximum performance. However, a benefit is that a cap may deter manipulation. In many jobs, extraordinarily high performance is unlikely to be possible, even for star employees, unless there was some manipulation (fixing the numbers, unethical or illegal conduct, etc.). This is another case where it is useful to keep in mind that a performance evaluation does not always measure performance well. When extremely high performance is implausible, except through manipulation, a cap is a good idea because it does not reward extreme performance.

Thus we see a cost and a benefit to having a cap. In many situations, both are relevant at the same time. Many jobs in banking, for example, involve star performers who can easily leave the firm if not compensated for very high performance, but manipulation is also possible and a serious concern. What can be done in such situations? Here, some subjectivity may help. A firm can use a discretionary cap, reserving the right to impose a cap on a bonus if it believes that very high performance may have resulted from either very good luck (uncontrollable risk that helped the employee's performance) or manipulation.

## CONCLUSIONS

This article has described how recent academic research, especially from the field of economics, provides a structured way to think about the design of an incentive plan. Two general conditions are important to the effectiveness of an incentive plan. Surprisingly, many reward schemes that firms use in practice violate one or both. The first is the effectiveness of the evaluation. It is critical that the evaluation reflects the employee's efforts, actions and decisions with reasonable accuracy and strength. Otherwise the evaluation will be beyond the control of the employee. Eventually the employee will figure this out (e.g., as the year-end bonus ends up having little to do with his/her actual work), and poor motivation will result. The effectiveness of the evaluation is by far the most important condition because a good evaluation method (that is, effective performance management) is not just fundamental to pay for performance, but also to employee feedback, training, communication and building of trust with the supervisor.

The second condition is that the evaluation has to be tied to rewards that are significant to the employee. That is not the same as saying that the reward is significant in magnitude. What is necessary for motivation is that the reward varies, and significantly, with employee performance. A large year-end bonus based on firm-wide profits will not be significantly affected by an individual contributor's efforts, no matter how hard the employee works. Instead, the employee will be rewarded for the efforts of everyone else as well, plus the effects of many uncontrollable factors. Once again, eventually the employee will figure out that whether he/she works hard or slacks off, the bonus check will still arrive in December. Again, many plans fail to meet this condition.

An incentive plan has two main components, which correspond to each of these conditions: the evaluation and how rewards are tied to the evaluation. A simple and useful structured approach to designing an incentive plan is to divide it into those two components and think about each issue separately.

The proper starting point is the evaluation. That is because the second step, tying rewards to the evaluation, depends on the properties of the evaluation method chosen. Furthermore, performance evaluation is far more complex

and causes most of the common problems with incentive systems. Therefore, it deserves the most careful thinking.

Evaluation should begin with an attempt to quantify performance with one or a small number of performance measures. In an ideal world, numeric measures that can be specified *ex ante* are preferred to subjective evaluation, because they give clear feedback to the employee, are more objective, and subjective evaluations are fraught with challenges.

Though performance measurement is complex and requires a fair amount of art, academic research helps by describing in relatively rigorous ways a key set of performance measure properties. These are distortion, uncontrollable and controllable risk, and manipulation. Most performance measures can be usefully analyzed with these concepts, since they arise in most situations. Thinking through them will also help the supervisor understand the employee's job better, which may lead to other improvements and better performance management. Each of these properties is important for thinking about the overall effectiveness of a performance measure. To the extent that the measure has serious flaws, it should be given less weight, or eliminated from, an incentive plan.

In addition, the flaws of numeric measures should be addressed through careful subjective evaluation. Subjective evaluation should largely be thought of fixing flaws in the quantitative component of the employee's evaluation. Subjectivity can be quite valuable in reducing distortions, ensuring the employee against uncontrollable risk while still holding that worker accountable for controllable risk, and detecting and deterring manipulation. A good subjective evaluation and performance management system has other benefits, of course, but the focus here is on how it can be used to improve the overall evaluation for incentive purposes.

The second step in designing an incentive system is to think about how rewards, both implicit and explicit, vary with the evaluation. The stronger the tie, the stronger the incentive. That is a good thing if the evaluation is effective, but it is why incentives will cause problems if the evaluation is ineffective. Therefore, the strength of the reward-evaluation tie depends fundamentally on the extent to which the evaluation provides balanced incentives, has limited measurement error, holds the employee accountable for manageable risks and is difficult to manipulate. Finally, the article also discussed related issues about the shape of the reward-evaluation relationship, including targets, lump-sum rewards, floors and caps.

Ultimately, incentive plan design is complex and requires substantial judgment. However, structured thinking with clear and rigorous principles will improve the design. Incentive compensation has been an extremely active area of research, both theoretical and empirical, in several academic fields. That accumulated body of research suggests a useful framework for just that kind of structured thinking, which can improve incentive plan design. ■

---

## AUTHOR

**Michael Gibbs, Ph.D. (michael.gibbs@chicagobooth.edu)** is a clinical professor of economics and faculty director of the executive MBA program at the University of Chicago Booth School of Business. He studies the economics of human resources and organizational design. He recently coauthored the second edition of *Personnel Economics in Practice*. Gibbs's research has been published in the *Quarterly Journal of Economics*, *Industrial & Labor Relations Review*, *Accounting Review* and other journals. He is a research fellow of the Center for the Study of Labor in Bonn, and of the Institute for Compensation Studies at Cornell University. In 2007, Gibbs received the Notable Contribution to Management Accounting Literature from the American Accounting Association. He has received two Hillel Einhorn Excellence in Teaching Awards from Chicago Booth. Gibbs earned a BA, MA and PhD in economics from the University of Chicago.

---

## REFERENCES

- Bushman, Robert M. and Abbie J. Smith. 2001. "Financial Accounting Information and Corporate Governance." *Journal of Accounting and Economics* 32: 237-333. <http://public.kenan-flagler.unc.edu/faculty/bushman/bushman-smith.pdf>. Viewed: Aug. 20, 2012.
- Courty, Pascal and Gerald Marschke. 2008. "A General Test for Distortions in Performance Measures." *Review of Economics and Statistics* 90(3): 428-441. <http://www.eui.eu/Personal/Courty/WkgPp/Courty%20and%20Marschke.pdf>. Viewed: Aug. 20, 2012.
- Gibbons, Robert. 1998. "Incentives in Organizations." *Journal of Economic Perspectives* 12(4): 115-132. <http://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.12.4.115>. Viewed: Aug. 20, 2012.
- Gibbs, Michael, Kenneth Merchant, Mark Vargus, and Wim Van der Stede. 2009. "Performance Measure Properties and Incentive System Design." *Industrial Relations* 48(2): 237-264. [http://www.cob.ohio-state.edu/~young\\_53/Gibbs-Merchant-VanderStede-Vargus.pdf](http://www.cob.ohio-state.edu/~young_53/Gibbs-Merchant-VanderStede-Vargus.pdf). Viewed: Aug. 19, 2012.
- Gibbs, Michael, Kenneth Merchant, Mark Vargus, and Wim Van der Stede. 2004. "Determinants and Effects of Subjectivity in Incentives." *The Accounting Review* 79(2): 409-436. [http://www.cob.ohio-state.edu/~young\\_53/Gibbs-Merchant-VanderStede-Vargus.pdf](http://www.cob.ohio-state.edu/~young_53/Gibbs-Merchant-VanderStede-Vargus.pdf). Viewed: Aug. 20, 2012.
- Healy, Paul. 1985. "The Effect of Bonus Schemes on Accounting Decisions." *Journal of Accounting and Economics* 7: 85-107. <http://www.sciencedirect.com/science/article/pii/0165410185900291>. Viewed: Aug. 19, 2012.
- Itnner, Christopher D. and David F. Larcker. 2002. "Determinants of Performance Measure Choices in Worker Incentive Plans." *Journal of Labor Economics* 20(2 pt. 2): S58-S90. <http://www.jstor.org/discover/10.1086/338674?uid=3739552&uid=2&uid=4&uid=3739256&sid=21101157238747>. Viewed: Aug. 19, 2012.
- Kohn, Alfie. 1975. "On the Folly of Rewarding A, While Hoping For B." *Academy of Management Journal* 18: 769-783.
- Murphy, Kevin J. 2000. "Performance Standards in Incentive Contracts." *Journal of Accounting and Economics* 30: 245-278. <http://www.sciencedirect.com/science/article/pii/S1573446399300249>. Viewed: Aug. 20, 2012.
- Murphy, Kevin J. 1999. "Executive Compensation." In *Handbook of Labor Economics*, ed. Orley Ashenfelter and David Card, 1st ed.1 vol.3: 2485-2563. New York: North Holland. <http://www.sciencedirect.com/science/article/pii/S1573446399300249>. Viewed: Aug. 20, 2012.
- Murphy, Kevin J. 1992. "Performance Measurement and Appraisal: Motivating Managers to Identify and Reward Performance." In *Performance Measurement, Evaluation, and Incentives*, ed. William J. Bruns. 38-51. Boston: Harvard Business School Press.
- Murphy, Kevin R. and J.N. Cleveland. 1995. *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives*. Thousand Oaks, CA: Sage Publications.
- Prendergast, Canice. 2002. "The Tenuous Trade-off between Risk and Incentives." *Journal of Political Economy* 110(5): 1071-1102. [http://www.nber.org/papers/w7815.pdf?new\\_window=1](http://www.nber.org/papers/w7815.pdf?new_window=1). Viewed: Aug. 20, 2012.
- Prendergast, Canice. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature* 37: 7-63. <http://bbs.cenet.org.cn/upload/22.202-01-4-6-12-35-32.pdf>. Viewed: Aug. 20, 2012.
- Prendergast, Canice and Robert Topel. 1993. "Discretion and Bias in Performance Evaluation." *European Economic Review* 37: 355-365. <http://faculty.chicagobooth.edu/canice.prendergast/research/DiscretionAndBiasInPerformanceEvaluation.pdf>. Viewed: Aug. 20, 2012.
- Roy, Donald. 1952. "Quota Restriction and Goldbricking in a Machine Shop." *American Journal of Sociology* 57(5): 427-442. <http://www.jstor.org/discover/10.2307/2772323?uid=3739552&uid=2129&uid=2&uid=70&uid=4&uid=3739256&sid=21101157238747>. Viewed: Aug. 20, 2012.