

MO MONEY, MO PROBLEMS: WHEN AND WHY FINANCIAL INCENTIVES BACKFIRE

Jennifer Burson, The University of Texas
Nolan Harvey, The University of Texas

ABSTRACT

This literature review examines when and why pecuniary incentives backfire. Contrary to what classical microeconomic theory might suggest, paying someone more might make them less willing to work on a task, charging a small fine might increase an undesired behavior and increasing the size of a reward might cause one's performance to deteriorate. In the last three decades, economists have made great progress in understanding incentives. However, much attention has been paid to a narrow and simplified view of human motivation while less research has aimed to understand why offering incentives can sometimes produce seemingly inconsistent outcomes. In this paper, we analyze relevant literature in behavioral economics and psychology as well as applications in various industries to understand under what circumstances monetary incentives can lead to diminished effort, motivation, or performance. We also aim to explain the underlying psychological phenomena that lead to these counterintuitive results.

INTRODUCTION

In the discipline of economics, incentives matter. *Freakonomics*, the bestselling book marrying economics and pop culture, goes so far as to define economics as “*the study of incentives*” (Levitt & Dubner, 2005). Extrinsic, monetary incentives like small cash payments, bonuses, or performance-related pay are common standbys for motivating or changing the behavior of individuals. An offer of a few dollars is the most common way to gain participants in academic research studies. Teacher's salaries are often tied to their student's performance on standardized tests. “*Paying for as*” is a common practice among parents in many middle and high schools. The list goes on; “*incentivizing*” a worker has become synonymous with offering payment.

Standard game theory takes for granted that “*raising monetary incentives increases [work] supply*” (Frey & Jegen, 2001) and “*a well-established result of most standard hidden action models is that higher incentives, ceteris paribus, lead to higher performance*” (Pokorny, 2006). In other words, the bigger and better the incentive, the more people will be motivated to perform a task (and therefore, their level of effort and performance will likely also increase in response to increased incentives).

However, it's easy to find exceptions to this seemingly obvious rule. Suppose someone asks their friend to help them move. Most friends would be willing to oblige. Now suppose the mover asks their friend to help them move in exchange for payment of a quarter. Economic theory tells us that the friend should be even more willing to help, since twenty-five cents is better than nothing—a “*higher incentive*” that should, in theory, “*lead to higher performance.*” In reality, though, it's hard to imagine that friend would be very pleased with the offer. Findings like these, where explicit incentives for task performance lead to decreased motivation and reduced long-run

performance, have been reported numerous times in psychology literature (e.g., Festinger & Carlsmith (1959), Deci (1972), Deci & Ryan (1985)).

This paper examines when and why pecuniary incentives backfire. In the last three decades, economists have made great progress in understanding incentives. However, much attention has been paid to a narrow and simplified view of human motivation while less research has aimed to understand why offering incentives can sometimes produce seemingly inconsistent outcomes.

Let it be noted that we cannot assume that monetary incentives simply do not work; in fact, there is a wide array of evidence supporting the basic premise of economics that incentives, particularly financial ones, are generally effective (e.g., Gibbons (1998), Prendergast (1999), Lazear (2000a,b) and Lee et al. (2017)). A more discriminating analysis is therefore required. This literature review aims to identify the circumstances under which monetary incentives can lead to decreased effort or worsened performance while exploring the underlying psychological phenomena that lead to these counterintuitive results.

Though there is a huge and largely uncharted territory of possible circumstances yielding unexpected results, we seek to understand only the most common circumstances which lead financial incentives to produce the opposite of the desired effect. These categories of common circumstances are: offering too little, paying when an agent is already intrinsically motivated, ignoring factors specific to the type of incentive (contingent or noncontingent) and designing or implementing an incentive in a problematic way. There is a certain amount of overlap in these sections; we categorize findings of previous research in this way in order to see trends within the existing literature to determine the most salient causes of failed incentives.

LITERATURE REVIEW

Paying Small or Not at All

Offering a very small amount of money as an incentive can lead to worse outcomes than offering no monetary incentive at all.

In Gneezy & Rustichini's (2000a) article "*Pay Enough or Don't Pay At All*," Israeli college students who were given a very small amount of money (the equivalent of three cents USD) for each IQ question answered correctly performed substantially worse than those who were offered no incentive for correct answers. As the student's IQ levels were controlled for (and the incentive was far too low to produce choking-see "*Choking*" section), we can safely assume the drop in performance to be due to decreased effort and motivation to perform.

In an experiment conducted by Heyman & Ariely (2004), hundreds of college students were asked to complete a brief survey. Within the survey, participants were asked about their willingness to load a sofa into a van after being given different types and levels of incentives. The students were offered three different forms of incentives: candy, cash, or "*monetized candy*," meaning the dollar value of the candy was mentioned. There were also three levels of incentives: a no-incentive control, \$0.50 or candy equivalent as the low-level incentive and \$5.00 or candy equivalent as the high-level incentive.

Heyman & Ariely (2004) proposed a dual market theory, which states that the relationship between effort and payments ultimately hinges on the type of exchange-whether it takes place within a money or social market. They describe a money market, in which transactions are highly sensitive to the size of compensation (in accordance with the standard model of labor, in which individuals trade their time and energy in exchange for monetary rewards) and a social market, in which effort is largely independent of compensation levels. Noncontingent incentives lean on this

social market and the obligation to reciprocate and contingent incentives depend heavily on the monetary market and its effects on output. (See more in “*Contingent vs. Noncontingent Incentives*” section.) In their view, the compensation (or lack thereof) determines the market in which the transaction takes place; each market has different underlying relationships between payment and effort.

Unpaid participants were assumed to be in a social market condition, whereas those paid cash (both in low and medium payment levels) were assumed to be in a money market condition. In the cases of students offered similar financial compensation in units of candy, Heyman and Ariely proposed that such exchanges would fall under a social market and thus expected different results from the scenarios involving monetary payment.

Their hypotheses proved correct; those incentivized with candy stayed within a social market transaction, where differences in size of pay (in this case, amount of unmonetized candy) did little to change the student’s willingness to help. The most telling outcome of this study is the comparison between willingness to help when nothing (or candy) was offered versus a low payment. Inconsistent with the typical economic view of humans-as-optimizers, participants were actually more willing to help when offered nothing than when offered 50 cents or a \$5 candy bar. This demonstrates that the 50 cents triggered a different interpretation for participants than the no payment treatment. Additionally, merely mentioning how much the candy was worth was enough to push the transaction back into the typical monetary market. These results, along with research by Pouliakas (2010), who concludes employers need to “*pay enough or don’t pay at all*”, point out the reality of counterproductive outcomes of offering contingent incentives that are too low (Figures 1 and 2).

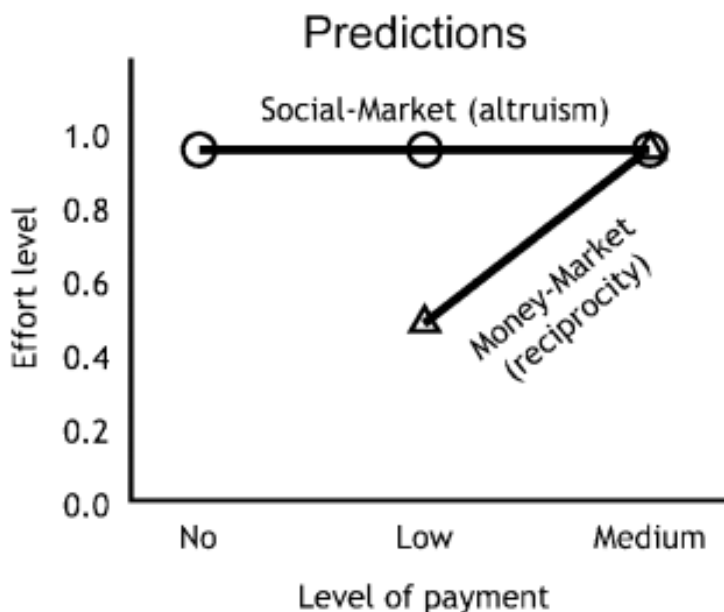


FIGURE 1
GRAPHICAL SUMMARY OF THE PREDICTED LEVELS OF EFFORT IN SOCIAL AND MONEY MARKETS (Heyman & Ariely, 2004).

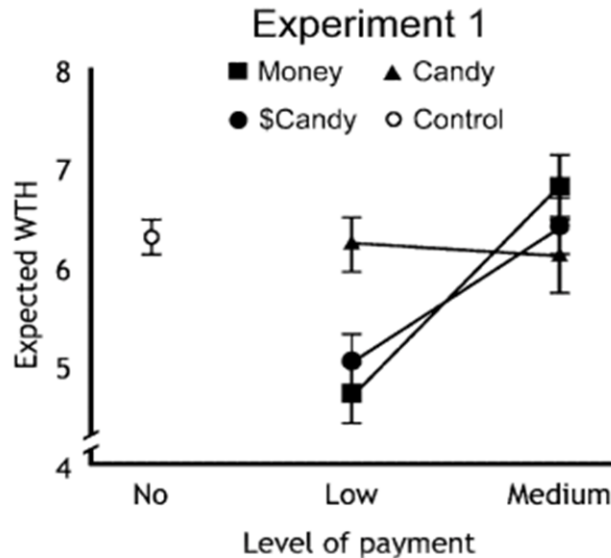


FIGURE 2

RESULTS FROM EXPERIMENT 1: EXPECTED WILLINGNESS TO HELP (WTH) AS A FUNCTION OF PAYMENT LEVEL (NONE, LOW, OR MEDIUM) AND PAYMENT FORM (MONEY, CANDY, OR MONETIZED CANDY) (Heyman & Ariely, 2004).

To demonstrate these effects in action (and not just through a self-report survey), Heyman and Ariely conducted two sets of experiments. In the first, participants were asked to drag an item on a computer monitor to a specific location on the screen. The participants exerted greater effort when paid \$4 as compared to 10 cents, but also exerted far less effort when paid 10 cents as compared to no payment at all. The results mirrored those demonstrated by the surveys: people were more willing to do a task for nothing than for a very small amount.

In the second experiment, participants were given a series of puzzles, the last of which was impossible to solve. Participants spent far less time trying to solve the last, unsolvable puzzle (pressing a give up' button earlier) if they were offered compensation than if they were given no incentive. It should be noted that this type of study design leaves some room for misinterpretation. One potential issue in the unsolvable puzzle study is that the amount of time participants spend on the unsolvable puzzle in condition x versus condition y is meant to stand in for effort. However, it's possible that those in condition x actually exerted more effort and realized sooner that the puzzle was unsolvable-thereby spending less time on the unsolvable puzzle than those in condition y. It is unclear if time spent on the puzzle is a clear stand-in for effort exerted. This potential issue is seen not only in Heyman and Ariely's study, but in many psychological experiments utilizing impossible tasks.

One interpretation of these results is that people find it demeaning to exchange their effort for very low sums. Ariely et al. (2008) had results consistent with this interpretation, finding that subjects demanded more money in exchange for their work building Lego structures if they witnessed their work being immediately disassembled by the experimenter (literally de-meaning their work).

Another, more complete explanation for why no incentives might be better than low incentives can be understood in the context of the principal-agent relationship. Pratt & Zeckhauser (1985) summarize the relationship thus:

“Whenever one individual depends on the action of another, an agency relationship arises. The individual taking the action is called the agent. The affected party is the principal.”

Any addition of money into the equation, even a small amount, has the potential to change the nature of the interaction between a principal and an agent. What was before viewed as a favor could be transformed in the mind of the agent as a market interaction when even fractions of a dollar are on the table (Fiske, 1991). Gneezy & Rustichini’s 2000b paper “*A Fine is a Price*” found that when a small fine was put in place for parents who were late picking up their children from daycare, the number of parents arriving late increased. While one might wonder if the fine was simply not large enough to generate the desired result, as *more* parents picked up their children late after the fine (as compared to the late-pickup rate remaining constant), the fine was not merely ineffective-it produced the opposite of the desired effect. This implies that the fine caused a change in perception toward tardiness. One interpretation of these findings is that the parents viewed the fine as a price for choosing the “*option*” of late pick-up: in other words, they felt justified in being late as long as they paid the small fine and expected no other negative consequences (like daycare worker’s disapproval). The fine transformed lateness from a rule violation that inconvenienced members of the daycare staff to a market transaction; in other words, the change in parent’s behavior was likely a result of a change in the transactional relationship between the parents and daycare.

Incentive Recipient Already Intrinsically Motivated

Another theme appearing in the literature is that of intrinsic versus extrinsic motivation. When an individual is extrinsically motivated, the common “*carrot or stick*” metaphor holds: the subject performs a desired behavior either for a reward (carrot) or to avoid a punishment (stick). Intrinsic motivation, on the other hand, does not require the external promise of reward or punishment-it exists within the individual and is typically sustained regardless of external conditions. Incentives of money, therefore, are inherently extrinsic motivators. Non-monetary rewards (such as candy, vacation, recognition, et cetera) can form the basis for either extrinsic or intrinsic motivation.

Nujjoo & Meyer (2012) studied the effects of certain incentives relative to the level of intrinsic motivation within individuals. They classified incentives into three types: extrinsic monetary, extrinsic non-monetary and intrinsic monetary (Figure 3).

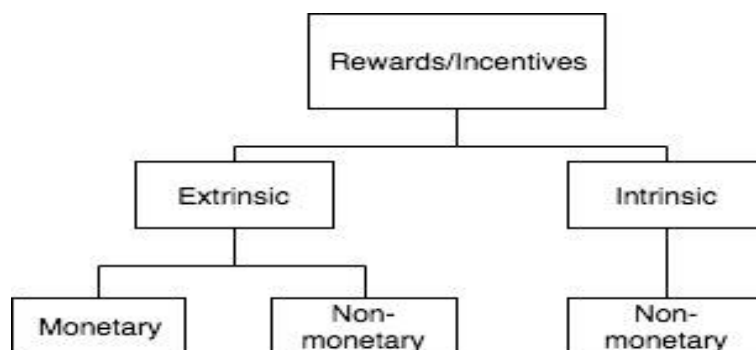


FIGURE 3
TYPES OF INCENTIVES CLASSIFIED BY EXTRINSIC OR INTRINSIC AND
MONETARY OR NON-MONETARY.

One of the primary explanations for why incentives can so often fail is the concept of “crowding out.” In macroeconomics, crowding out refers to government spending decreasing private company’s motivation for investing in infrastructure or public good projects. In education and psychology, where student’s intrinsic versus extrinsic motivations are of great concern, this concept often refers to material rewards for academic performance harming a student’s internal desire to learn or master a subject. In examining why offering money for altruistic acts might backfire, extrinsic motivators (in the form of monetary incentives) might “crowd out” one’s intrinsic motivations to perform an altruistic act. In other words, if the primary motivator for a task is internal and not based on some external reward, offering the external reward might, in the mind of the agent, overpower their intrinsic motivation.

On the other side of the coin, Akerlof & Dickens (1982) suggest that imposing stiffer penalties for some crimes might sometimes be counterproductive. When consequences are negligible, those who obey laws and rules feel the psychological drive to create internal justification for their actions (since intellectually, the minimal sanction shouldn’t really be enough to change their behavior). This means that they are inhibited by their own justification and therefore less likely to break the law or rule, even when the threat of punishment is low. Larger penalties have the potential to undermine these internal justifications, so once the threat of the penalty no longer exists, lawbreaking increases. Intensifying the consequences could, in actuality, crowd out individual’s moral compasses. This “crowding out” has also been viewed in the realms of social interactions, provision of public goods, tax compliance, volunteering and experimental labor contracts. (Frey, 1997; Frey & Jegen, 2001). Frey and Jegen also view this crowding out effect as a negative shift of a supply curve, relating effort exerted and the value of an extrinsic incentive (Figures 4 and 5).

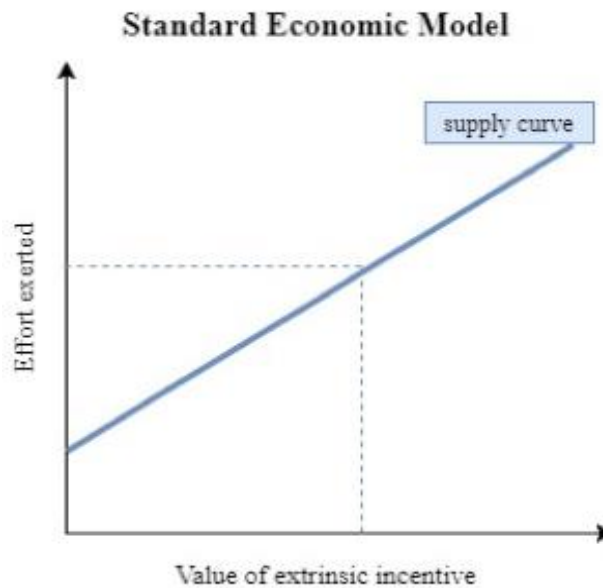


FIGURE 4
STANDARD SUPPLY CURVE SHOWING EFFORT EXERTED AS A FUNCTION OF THE VALUE OF AN EXTRINSIC INCENTIVE.

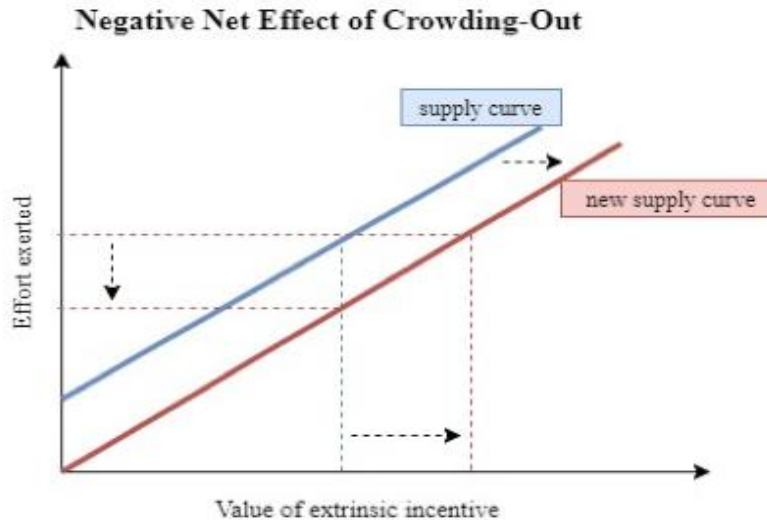


FIGURE 5
A NEGATIVE SHIFT IN THE SUPPLY CURVE AS A RESULT OF CROWDING OUT.

Another possible explanation is that effort decreases due to the “*overjustification effect*.” Similar to the “*crowding out*” concept, this effect occurs when a task or behavior is already satisfying (or at least, not unpleasant) to an agent. Thus, offering a payment can imply that a task is unpleasant, leading to a decreased willingness to perform on the part of the agent relative to their previous positive feelings about the task. This result was demonstrated by Lepper et al. (1973), who found that people’s intrinsic interest in a target activity is decreased when they are extrinsically motivated to meet a goal through said target activity.

The results of crowding out and the overjustification effect are both consequences of cognitive dissonance and our subsequent drive to correct dissonance and restore consistency. Cognitive dissonance occurs when contradictory attitudes, beliefs, or behaviors produce mental discomfort, leading a person to alter one of the attitudes, beliefs, or behaviors to reduce the discomfort and restore psychological consistency (Festinger, 1957). Using external cues to shape our internal drive is done every time we assume a more expensive bottle of wine is inherently better. Likewise, underperforming due to a proffered incentive is a direct result of our unconscious drive to reduce cognitive dissonance.

Prosocial behaviors

Prosocial behaviors, commonly defined as those acts which are costly to oneself and beneficial to others, are an interesting subset of potentially incentivized behaviors to which people often react paradoxically. Because motives for performing prosocial behaviors are often intrinsic-focusing either on the opinion of others or one’s own self-image-offering a financial reward for doing such altruistic acts often undermines the agent’s true motive.

Gneezy & Rustichini (2000b) found that schoolchildren collecting donations for a charitable organization collected less money when given performance incentives. Similarly, Frey & Götte (1999) found that volunteers, when rewarded with pay, reduced the number of hours they volunteered to work. Titmus (1970) argued that offering payment for blood donations might decrease willingness to give blood. To test this assertion, Mellstrom & Johannesson (2008) conducted a field experiment comparing willingness to donate when subjects were compensated,

not compensated, or given an option to donate the payment to charity. None of these three treatments altered the willingness to donate among male subjects, but interestingly, offering payment halved the number of female donors.

One interpretation of results such as these is that people are motivated to perform prosocial behaviors, at least in part, by wanting to gain or maintain the goodwill of others. Bénabou & Tirole (2003) explain that social pressure and norms that attach honor to good deeds and shame to selfish ones can be a primary motivator for helping or altruistic acts. Nonprofits often take advantage of individual's desire to publicly display generosity (or at least the appearance thereof) through donation recognitions from lapel pins to t-shirts to plaques on benches. These "*social signaling*" motives for giving are perhaps most evident by the extreme rarity of anonymous donations: donations without an attached name (which have the exact same tax-deduction benefits as those attributed to someone) comprise less than one percent of all donations in any given year (Glazer & Konrad, 1996).

A second interpretation is based on the notion that prosocial behaviors are sought out and performed due to the desire to enhance one's own self-image—that is, being able to "*look at oneself in the mirror*" and, as if through the eyes of another, like the person we see. This understanding of self-image-enhancing motivations is nothing new; Smith (1776) described altruistic motivations in terms of individuals judging their own conduct and morals through the eyes of an "*impartial spectator*," an "*ideal mate within the breast*." Batson (1998) writes that "*the ability to pat oneself on the back and feeling good about being a kind, caring person, can be a powerful incentive to help*." Batson also discusses the anticipation of guilt and the desire to avoid such self-image-damaging experiences. With these types of motivations in mind, it's easy to understand that providing money or other visible incentives for being altruistic may lead people to "*interpret their motivation as egoistic even when it is not*," thus negating their true reason for performing such generous acts (Batson, 1998). Social signaling motives are inherently intrinsic; as Frey & Jegen (2001) describe

"An intrinsically motivated person is deprived of the chance of displaying his or her own interest and involvement in an activity when someone else offers a reward, or orders him/her to do it."

Contingent vs. Noncontingent Incentives

We can broadly classify incentives into two categories: contingent and noncontingent. Contingent incentives are rewards that are received upon the successful completion of a certain task or benchmark. Noncontingent incentives are given to a person regardless of their future performance. The success of noncontingent incentives is based largely on people acting on their propensity to reciprocate. Both contingent and noncontingent incentives are used in a variety of ways to motivate employees and generate higher output.

These two types of incentives can be used exclusively or in conjunction with one another. In most cases, organizations are confident in implementing one of the two incentives and they are faced with the decision of how large to make the incentive or to whom it will be offered. However, in some cases, decisions must be made between offering contingent or noncontingent rewards. In Gneezy & Rey-Biel's 2011 study, "*On the Relative Efficiency of Performance Pay and Social Incentives*," experimenters sent out letters asking customers to fill out a survey. Some letters contained an upfront payment ranging from \$1 to \$30, while others promised payment upon completion of the survey. It was discovered that the upfront payment generated a boost in the response rate and the size of the payment seemed to have a relatively minor effect. This finding

pointed to the potential issue of overpaying when using noncontingent incentives. Regarding the letters that promised payment upon compliance, the response rate was low for the small offerings, but increased as the promised reward increased. This result exhibited that the size of the contingent incentive is a key factor in determining the ultimate success of its implementation. Additionally, they found that upfront payments were consistently more costly per response than the contingent payments.

Contingent incentives

Contingent incentives have the potential to fail if some factors are not adjusted carefully. Most notably is the size of the incentive. Contingent financial incentives can backfire if they are too low or even too high, resulting in an effort level that is lower than the baseline.

When considering contingent incentives, there is a monetary threshold that must be exceeded-otherwise; the low incentive can result in lower output compared to a no-incentive control setting. Gneezy & Rey-Biel (2011) discovered in their aforementioned study on survey response rates that there was a minimum dollar amount that needed to be promised to generate a higher response rate than the no-incentive control. The control (no incentive offered) response rate to the survey was 7.6%. When offering a future payment of any amount less than \$6, the response rate was lower than the control. Since performance-contingent incentive compensation systems are cited as key elements of any high-performance work system (HPWS), it is crucial that such incentives be designed to produce the desired results (Arik & Geho, 2017). The offering of a contingent incentive that is too small can be counterproductive in a way that results in less effort than would be observed if there was no incentive at all. One potential explanation for this finding is when subjects associate a monetary value to a task, they will then judge whether it is worthy or not to input the required effort and time.

Choking

Research suggests that the contingent incentives offered can also be too high, causing a negative output compared to a no-incentive or previously incentivized environment. Baumeister (1984) found that situations where perceived pressure to perform is too high can yield worsened results. In other words, subjects will perform at a lower level when facing a higher-value incentive (this worsened performance due to increased pressure or heightened nervousness is colloquially known as “choking”). Baumeister’s subjects performed a “*practice run*” of a trivial task then were offered a cash prize if they met a goal that far exceeded their practice run. Baumeister observed that on the first trial after the practice run, many subjects choked, performing significantly worse than they had in the practice run.

Expanding on the existing research, Ariely et al. (2009) examined factors that contribute to a subject choking such as the environment, personal characteristics and the type of task being performed. In their first experiment, residents in a rural town in India participated in six different games. Some games tested motor skills while others tested cognitive ability and memory. There were low, medium and high payouts (defined by the relative equivalent in participant’s salaries) to incentivize the subjects. The subjects consistently underperformed in all games when offered the high-value incentive, demonstrating the effects of choking when the stakes are too high. This mirrors the early-1900s “*Yerkes-Dodson Law*,” which states that performance improves with an increase in physiological or mental arousal-but only up to a certain point, after which performance worsens (Yerkes, 1908).

In a second experiment, Ariely et al. found that the magnitude of these choking effects differ based on the type of task being performed. Subjects performed two types of tasks in the face of varying levels of incentives. The first task was quickly pressing two keys on a computer keyboard (a physical task) and the second was finding two numbers that added up to ten in a matrix of twelve numbers (a cognitive task). When high-value payouts were offered, choking effects were present in both tasks; the cognitive task, however, yielded much stronger choking effects, suggesting that the pressure of the high-value payout impeded on subject's cognitive abilities more than their physical abilities. Essl & Jaussi (2017) found similar results when they added unrealistic time constraints to tasks: time pressure has been shown to “*inhibit logic-based reasoning*” needed for such cognitive tasks (Leaptrott & Mc-Donald, 2008).

Oddly, the results showed that those subjects who choked on one task did not always choke on the other; in fact, there was a negative correlation, showing if an individual choked on one type of task, they would be less likely to choke on the other type of task. This suggests a possible interaction between individual characteristics, type of task and level of incentive. More research is necessary to isolate the individual differences that could have led to this finding.

Noncontingent incentives

Noncontingent incentives are built upon the social principle of reciprocation and set the premise for a “*gift exchange*.” To fully understand the changes in behavior in response to incentives, we need to recognize motives like the desire to reciprocate (Fehr & Falk, 2002). This motive to reciprocate can be affected by innate characteristics or external perceptions of an environment. For example, though an individual's salary is indirectly linked with their performance, the relationship cannot be said to be direct (as it would be if a portion of your annual pay was doled out only upon completion of each task you performed throughout a work day). Thus, when employers pay a salary, they are counting on workers to return the “*favor*” with a certain level of effort. Camerer et al. (2004) frame the employer-employee relationship in terms of pure, reciprocation-based principles: employers pay employees as a gift in the hopes the workers repay the gift by working harder than is required. This inherent desire to reciprocate is always present in individuals, but can vary in how much it influences individual's actions (Fehr & Falk, 2002; Cohn et al., 2014). There are some people who simply leave their performance unchanged in the face of an incentive; Cohn et al. (2014) classify these individuals as “*nonreciprocal*.” In the case of a nonreciprocal individual, noncontingent incentives will likely be ineffective. Parnell & Sullivan (1992) propose that the effectiveness of such performance-based pay systems is a direct function of a given individual's “*equity sensitivity*”; those who prefer to be equitably compensated (rather than over-or under-paid) are more likely to find pay-for-performance systems motivating and satisfying.

In addition to the variable magnitude of reciprocation motives among individuals, environmental factors can lead this reciprocal relationship to break down, such as when the environment is perceived as unfair, there is no sense of community, or the employee does not have a good relationship with his/her supervisor (Nelson, 2017). A large factor in determining the success of a noncontingent financial incentive is the perception of fairness held by the individual who receives it. In short, if the working environment is not perceived as fair, the individual will be less likely to reciprocate. Cohn et al. (2014) conducted a field experiment in which they compared the effects of a pay raise between employees who perceive being underpaid versus adequately paid at the base wage rate. This experiment found that when the worker already perceived the environment as fair (as they perceived their pre-incentive pay as adequate), they rarely produced

more output. This observation yielded a key finding that noncontingent incentives in the form of a pay raise act to *remove negative reciprocity* in a work environment (or eliminate some perceived unfairness), rather than adding positive reciprocity. This means that the noncontingent incentive can fail to produce increases in performance or output if the environment is already perceived as fair.

Implementation

This paper discusses a variety of factors that yield counterproductive results following the introduction of financial incentives. Many of these factors are counterintuitive, running against the classical economic theory, which states that higher incentives yield higher output. An incentive structure can motivate the correct inputs, be the right size and fit all other criteria for an effective financial incentive-but still backfire due to the way the system is implemented. Like any system, if not implemented correctly, the incentive can fail (or not be as effective as possible) regardless of its design.

Although strategic decision-making (like implementing an incentive system meant to boost productivity or performance) is often seen as an objective exercise (especially when supported by fancy, expensive analytical tools), a kernel of common sense, viewpoint-taking and empirical research results should be the basis of any decision made that affects human beings (i.e. all decisions) (Borrero & Henao, 2017). Any time a financial incentive is being considered, managers need to analyze whether the incentive is implemented in a way that successfully generates additional output to the benefit of the firm. An organization should be cautious before implementing an incentive and should ensure that the right actions are being incentivized. Baker et al. (1988) point out those piece-rate workers may not perform their job in a way that ultimately benefits the employer. Although individual differences certainly mediate responses to quantity-based pay, this type of payment system results in many workers sacrificing quality for quantity. Similarly damaging, managers who offer incentives for annual accounting profits can potentially sacrifice long-term profitability for short-term earnings. Additionally, worker's motivation and effort input can be adversely affected due to mismanagement of pay-per-performance programs (Hamner, 1975).

Offering financial incentives to employees is an act that is not forgotten and can change the baseline expectation of employees. If a worker is given a bonus one year, then given a smaller bonus or no bonus at all the following year, it can yield a lower output overall relative to the smaller bonus given both years. Because of this, managers should be cautious before initially implementing an incentive, revoking an incentive from one year to the next, or changing objective performance measures (Baker et al., 1988; Georgellis & Tabvuma, 2010; Pouliakas, 2010).

In a 2015 study by Angelucci et al. it was discovered that the amount of people who publicly receive the same incentive can affect the intended result. Angelucci et al. studied schoolchildren and their reactions to fluctuating levels of incentives for choosing grapes as a snack over a cookie. In the study, the percentage of students who received the incentive varied, along with whether the incentive was public or private. One focus of the research was that some children may conform to others who picked the grapes (even if those who initially picked the healthy snack were incentivized to do so). However, when children discovered that every other student received the incentive to choose the grape as well, the reactions to the incentive were consistently negative. Whenever the incentive was made public, students who did not receive the incentive were less likely to conform to the choice of the grapes; even when all students were incentivized publicly, there was a low take-up of the grapes option. The public awareness of an incentive can result in

negative outputs, create competition amongst co-workers instead of teamwork and strain the relationship between workers and their superior (Kohn, 1993).

Suggestions for Future Research

The field of behavioral economics is still relatively new, meaning that there are lots of gaps and opportunities in the literature. In our free-enterprise economy, incentives are commonly used as shortcuts for providing motivation. Properly designed incentives can be a potent way to induce desired behaviors, but poorly constructed incentives can be not only ineffective, but counterproductive. This topic of research is pertinent to nearly every industry: further investigation is needed in order for empirically-based recommendations to be made.

One of the most immediate needs for empirical findings is in cases where incentive systems are being used or considered, but the full effect of their implementation has perhaps not yet been considered. One example of such a case is organ donations. There is a global shortage of organs available for transplantations; this huge gap between supply and demand of human tissues has led to overseas transplantations often in developing countries, black markets for organs and the illegal, often dangerous, medical procedures that accompany such black-market transactions (Adair & Wigmore, 2011). Some countries have banned paid donations altogether, while other countries have experimented with paying for donations of tissues from living donors (for example, kidneys and eggs). By “*transplanting*” organs into a money market, people will almost certainly strive for a profit margin, even if this results in human exploitation. Additionally, the sale of human tissues means, at a basic level, that life-saving organs go to those rich enough to afford them while “*donations*” will come from those who need the money. It’s hard not to see how a system of this kind would perpetuate inequality and raise major ethical concerns.

Aside from the ethical implications of putting a price tag on human tissues, the effectiveness of such a system should be of great concern before its implementation. Paying for organs puts a price tag on the act of donation, so a potential kidney donor, for instance, who previously thought “*I give a kidney, someone gets to live*” could interpret the trade as “*I give a kidney in exchange for x amount.*” Such a switch in interpretation could potentially decrease donations (Adair & Wigmore, 2011).

While this paper deals with financial incentives backfiring—that is, producing the opposite result to the one desired—incentives might also lead to unintended consequences if they are not designed to incentivize the proper inputs in addition to producing the desired result. This is a particularly relevant issue and one that requires much greater attention given the widespread corruption, abuses of power and mismanagement seen in headlines daily (Covelli & Mason, 2017). One recent example demonstrating the consequences of these misaligned incentives is Wells Fargo’s 2017 fraud scandal. By setting unrealistically high sales targets and account quotas above what was possible, managers and sales representatives were incentivized not to keep their customers happy or convert customers from competing banks, but to open as many accounts as possible. Through misaligned incentives (encouraging increased number of accounts rather than existing customer profitability or fee income) and excessively high financial targets, Wells Fargo created an environment in which its employees often had (or felt like they had) little choice but to cheat the system. A huge fraud scandal could have been avoided if top management was mindful of exactly what actions they were incentivizing through their aggressive goals.

Another area lacking in research is the degree to which the type and sizes of incentives exist on a continuum. In other words, what type of relationship exists between the size of the reward and effectiveness of the incentive: linear, u-shaped, exponential, or other? Additionally, is

there a cutoff point at which most people agree a reward is “*too small*” or “*too large*”, or do individual interpretations of the smallness or largeness of an incentive depend on age, economic background, current salary level, personality traits, or culture?

This paper focuses specifically on incentives in the form of fiat money in order to normalize the value of the incentive across interpretations-some research has looked at non-monetary incentives (like candy) and if the same results hold true when those non-monetary rewards are monetized (e.g. offering a “*50 cent candy bar*” vs. a candy bar). More work must be done on types of nonmonetary incentives, as these can be powerful motivators and in some cases, might not provide the same challenges that lead financial incentives to backfire. Nonmonetary, social rewards (such as public recognition) in particular would likely not impinge on intrinsic motivation in the same way. Individual differences in acceptance of and motivation for nonmonetary incentives is also a relatively untouched area of research that deserves more attention. However, it seems that social rewards can easily be turned into extrinsic rewards by the mere mention of value (“*I will buy you lunch*” as compared to “*I will buy you a \$15 lunch*”), so the characteristics of social rewards deserves further investigation.

Presentation or media factors and their impact on the results deserve further attention as well. Do people react the same way to *seeing* an incentive of \$30 as they do to *hearing* of a thirty dollar incentive? Does the presence of the experimenter mediate the effects of accepting incentives for prosocial behaviors? Opportunities exist here to measure how much someone’s motivation to participate in prosocial behaviors is based on self-image and how much is based on the opinion of others. This matters because small tweaks (like signing up online while alone versus on a public signup sheet) could make all the difference in whether incentives motivate or demotivate.

One study (Mellstrom & Johannesson, 2008) found a significant difference in the effect of offering compensation for blood donations between male and female participants. This suggests more research is needed to adequately determine whether different schemes should be considered for different genders and other differing demographics.

As mentioned earlier, study design can potentially lead to misinterpretation of results (as in Heyman & Ariely’s, 2004 study using an unsolvable puzzle). It should be noted that effort and performance are measured in a variety of ways across studies referenced in this paper. These study designs could be improved by ensuring that the measures of effort or performance closely line up with the study’s most salient application. Some studies measure effort and others, performance-none aim to measure both, though, which would be an important data point in demonstrating instances like choking (in which effort might be high but performance low) or “*impossible task*” performance (in which the measure of performance-time spent trying to solve an impossible task does not necessarily align with exerted effort).

CONCLUSION

In Mark Twain’s classic 1876 novel, the eponymous Tom Sawyer tricks a group of boys into doing his chores for him by making them think he enjoys painting fences (a task he actually finds loathsome). He even gets them to pay him for the “*privilege*” of painting, concluding that “*Work consists of whatever a body is obliged to do and play consists of whatever a body is not obliged to do.*” Twain further remarks,

“There are wealthy gentlemen in England who drive four-horse passenger-coaches twenty or thirty miles on a daily line in the summer because the privilege costs them considerable money; but if they were offered wages for the service, that would turn it into work and then they would resign.”

Twain knew long ago what many managers, leaders and other motivators today overlook. Human beings have much more flexible perceptions about what they are willing to pay for and what they're willing to do for payment than game theory would suggest; cognitive psychology shows us many of the shortcuts-heuristics and cognitive biases, for example-people use to derive value while saving mental energy. Humans are not programmed to always respond in ways which maximize their output, monetary or otherwise: if they were, we should be very confused by the sheer number of illogical choices made every day.

This paper outlines the most prominent themes in the literature regarding what pitfalls to watch out for that could cause an incentive to produce a counterproductive result - i.e., backfire. There is still much work to be done in establishing which incentives work in what circumstances; in the meantime, those wishing to motivate through monetary incentives should not assume that financial incentives will necessarily motivate all people or all tasks and should monitor the impact of incentives after implementation for unintended consequences or even worsened results. More attention must be paid to the factors detailed in this paper-the level and type of incentive should be aligned with the recipient's motive, the relationship between principal and agent should be maintained and the implementation of the incentive should be carefully done-lest the intended carrot become, in actuality, a stick.

In short, applying game theory takeaways broadly can undermine our objectives when human factors are not considered. If we fail to recognize that we're aiming to motivate and incentivize human beings and not computers, we will ignore many (and often the most significant) factors in designing and implementing these systems. When it comes to motivating people, bringing in mo' money might mean mo' problems.

REFERENCES

- Adair, A., & Wigmore, S.J. (2011). Paid organ donation: the case against. *Annals of The Royal College of Surgeons of England*, 93(3), 191-192.
- Akerlof, G.A., & Dickens, W.T. (1982). The economic consequences of cognitive dissonance. *American Economic Review*, 72(1), 307-319.
- Angelucci, M., Prina, S., Royer, H., & Samek, A. (2015). When incentives backfire: Spillover effects in food choice. *National Bureau of Economic Research*.
- Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. (2009). Large Stakes and Big Mistakes. *Review of Economic Studies*, 76(2), 451-469.
- Ariely, D., Kamenica, E., & Prelec, D. (2008). Man's search for meaning: the case of Legos. *Journal of Economic Behavior & Organization*, 67(3), 671-677.
- Arik, M., & Geho, P. (2017). Skill gaps, human capital formation, and strategic choices: What do businesses do when facing critical labor force challenges? *Global Journal of Management and Marketing*, 1(2), 111-125.
- Baker, G.P., Jensen, M.C., & Murphy, K.J. (1988). Compensation and incentives: practice vs. theory. *The Journal of Finance*, 43(3), 593-616.
- Batson, D. (1998). Altruism and prosocial behavior. *Handbook of Social Psychology*, 2(1), 282-316.
- Baumeister, R., & Hogan, R. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, 46(3), 610-620.
- Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70(244), 489-492.
- Borrero, S., & Henao, F. (2017). Can managers be really objective? Bias in multicriteria decision analysis. *Academy of Strategic Management Journal*, 16(1), 11-20.
- Camerer, C., Loewenstein, G., & Rabin, M. (2011). *Advances in behavioral economics*. Princeton, NJ: Princeton University Press.
- Cohn, A., Fehr, E., & Goette, L. (2014). Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. *Management Science*, 61(8), 1777-1794.
- Covelli, B.J., & Mason, I. (2017). Linking theory to practice: Authentic leadership. *Academy of Strategic Management Journal*, 16(3), 20-34.

- Deci, E.L. (1972). Intrinsic motivation, extrinsic reinforcement and inequity. *Journal of Personality and Social Psychology*, 22(3), 113-120.
- Deci, E., & Ryan, R.M. (1985). Intrinsic motivation and self-determination in human behavior. *Perspectives in Social Psychology*. New York: Plenum.
- Essl, A. & Jaussi, S. (2017). Choking under time pressure: The influence of deadline-dependent bonus and malus incentive schemes on performance. *Journal of Economic Behavior and Organization*, 133(1), 127-137.
- Fehr, E., & Falk, A. (2002). Psychological foundations of incentives. *European Economic Review*, 46(4), 687-724.
- Festinger, L. (1962). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Festinger, L., & Carlsmith, J. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal Psychology*, 58(2), 203-210.
- Fiske, A.P. (1991). *Structures of social life: the four elementary forms of human relations: communal sharing, authority ranking, equality matching, market pricing*. New York: Free Press.
- Frey, B. (1997). A constitution for knaves crowds out civic virtues. *Economic Journal*, 107(443), 1043-1053.
- Frey, B.S., & Goette, L. (1999). *Does pay motivate volunteers?* Institute for Empirical Research in Economics, University of Zurich.
- Frey, B.S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, 15(5), 589-611.
- Georgellis, Y., & Tabvuma, V. (2010). Does Public Service Motivation Adapt? *Kyklos*, 63(2), 176-191.
- Gibbons, R. (1998). Incentives in organizations. *The Journal of Economic Perspectives*, 12(4), 115-135.
- Glazer, A., & Konrad, K. (1996). A signaling explanation for charity. *The American Economic Review*, 86(4), 1019-1028.
- Gneezy, U., & Rey-Biel, P. (2014). On the relative efficiency of performance pay and noncontingent incentives. *Journal of the European Economic Association*, 12(1), 62-72.
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *The Journal of Legal Studies*, 29(1), 1-17.
- Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all: the effect of monetary compensation on performance. *Quarterly Journal of Economics*, 115(3), 791-810.
- Hamner, W. (1975). How to ruin motivation with pay. *Compensation & Benefits Review*, 7(3), 17-27.
- Heyman, J., & Ariely, D. (2004). Effort for Payment: A Tale of Two Markets. *Psychological Science*, 15(11), 787-793.
- Kohn, A. (1993). Why incentive plans cannot work. *Harvard Business Review*, 71(5), 54-62.
- Lazear, E. (2000a). Performance pay and productivity. *The American Economic Review*, 90(5), 1346-1361.
- Lazear, E. (2000b). The power of incentives. *The American Economic Review*, 90(2), 410-414.
- Leaptrott, J., & Mc-Donald, J.M. (2008). O Assessing managerial decisions using the dual systems theory of reasoning: Future challenges for management researchers. *Academy of Strategic Management Journal*, 7(2), 25-33.
- Lee, C.C., Strohl, K., Fortenberry, M., & Cho, Y.S. (2017). Impacts of human resources management innovations on productivity and effectiveness in a medium-size non-profit organization. *Global Journal of Management and Marketing*, 1(1), 36-55.
- Lepper, M.R., Greene, D., & Nisbett, R.E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28(1), 129-165.
- Levitt, S.D., & Dubner, S.J. (2005). *Freakonomics: A rogue economist explores the hidden side of everything*. New York: William Morrow.
- Mellström, C., & Johannesson, M. (2008). Crowding out in blood donation: Was titmus right? *Journal of the European Economic Association*, 6(4), 845-863.
- Nelson, T.A. (2017). LMX relationships and "social death": A curvilinear effect on ostracism. *Global Journal of Management and Marketing*, 1(1), 55-78.
- Nujjoo, A., & Meyer, I. (2012). The relative importance of different types of rewards for employee motivation and commitment in South Africa. *SA Journal Of Human Resource Management*, 10(2), 111-124.
- Parnell, J.A., & Sullivan, S.E. (1992). When money isn't enough: The effect of equity sensitivity on performance-based pay systems. *Human Resource Management Review*, 2(2), 143-155.
- Pokorny, K. (2008). Pay-but do not pay too much: An experimental study on the impact of incentives. *Journal of Economic Behavior & Organization*, 66(2), 251-264.
- Pouliakas, K. (2010). Pay enough, don't pay too much or don't pay at all? The impact of bonus intensity on job satisfaction. *Kyklos*, 63(4), 597-626.
- Pratt, J., & Zeckhauser, R. (1985). Principals and agents: the structure of business. *Harvard Business School Press*.
- Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature*, 37(1), 7-63.
- Smith, A. (1759). *The theory of moral sentiments*. London: Printed for A. Millar, and A. Kincaid and J. Bell.

- Titmus, R.M. (1970). *The Gift Relationship*. London: Allen and Unwin.
- Twain, M. (1876). *The Adventures of Tom Sawyer*. New York: Library of America.
- Wallace, C., Combs, S., Jordan, S., & Betha, M. (1997). Mo Money Mo Problems [Recorded by B. Smalls]. *On Life After Death* [CD]. Steven Jordan. (1996)
- Yerkes, R.M., & Dodson, J.D. (1908). *The relation of strength of stimulus to rapidity of habit-formation*. Granville, OH: Editorial Office, Denison University.