THE DARK SIDE OF GOAL SETTING: THE ROLE OF GOALS IN MOTIVATING UNETHICAL DECISION MAKING

MAURICE E. SCHWEITZER
Wharton School
University of Pennsylvania
Philadelphia, PA 19104

LISA ORDÓÑEZ AND BAMBI DOUMA
University of Arizona

ABSTRACT

This article demonstrates that goal setting motivates unethical behavior. This is true for goals both with and without economic incentives, and we find that the relationship between goal setting and unethical behavior is particularly strong when people are just short of reaching the goal. We explain our results in terms of the reference point adoption process and Prospect Theory (Heath, Larrick, & Wu, 1999; Kahneman & Tversky, 1979).

INTRODUCTION

A substantial literature has documented the benefits of setting goals. In general, people exert more effort and work more persistently to attain difficult goals than they do when they attempt to attain less difficult goals or "do their best" (Locke & Latham, 1990). This relationship is so strong that goal setting has become an important part of motivation theory and management education. In practice, goal setting has been recommended as a practical approach to enhance work performance (Ivancevich, 1974), and Locke and Latham (1990) contend that goal setting may be the most effective managerial tool available.

In this article, we argue that goal setting should be used cautiously. We extend prior work by demonstrating that difficult goals can motivate not only constructive behavior, but also unethical behavior. We conceptualize goals as reference points (Heath, Larrick, & Wu, 1999) and use the Prospect Theory value function (Kahneman & Tversky, 1979) to develop specific predictions relating goal setting and unethical behavior.

Unlike standard theories of economic utility, Prospect Theory describes preferences as labile and referent dependent—outcomes are judged relative to a specific reference point. According to Prospect Theory outcomes better than the reference point are evaluated as gains, and outcomes worse than the reference point are evaluated as losses. Consistent with Heath, Larrick, and Wu (1999) we consider goals as potential reference points, and we expect people with goals to evaluate outcomes short of the goal as losses, and outcomes greater than the goal as gains.

The Prospect Theory value function is different over the domain of gains than it is over the domain of losses. Specifically, the value function is (a) steeper in the domain of losses than it is in the domain of gains and (b) convex in the domain of losses and concave in the domain of gains. The change in slope implies loss aversion; changes in outcomes in the loss domain are

valued more than equivalent changes in the gain domain. For example, Prospect Theory implies that people value a $100 loss more than they value a $100 gain. With respect to goal setting, this change in valuation implies that an individual may be more motivated to increase performance prior to reaching a goal than he or she is to increase performance after reaching a goal. This motivation may extend to both ethical behavior and unethical behavior. Specifically, we consider the following related hypotheses that link goal setting and unethical behavior:

> *Hypothesis 1: Decision makers who are short of reaching a goal will be more likely to engage in unethical behavior than decision makers who do not have a set goal.*

> *Hypothesis 2: Decision makers who are short of reaching a goal will be more likely to engage in unethical behavior than decision makers who exceed the goal.*

The change from convexity to concavity in the Prospect Theory value function implies that changes in outcomes just short of the reference point will be valued most highly. Prospect Theory implies that people will value reducing a loss by $100 <u>to break even</u> more than they value reducing a loss from $2,000 to $1,900. With respect to goal setting, this change in valuation implies that an individual may be more motivated to increase performance in the loss domain <u>near</u> a goal than he or she is to increase performance in the loss domain <u>far from</u> a goal. As before, we expect this motivation to extend to both ethical and unethical behavior.

> *Hypothesis 3: Decision makers who are close to reaching the goal will be more likely to engage in unethical behavior than similar individuals who are far from reaching the goal.*

SCENARIO STUDIES

We conducted eight scenario studies and an anagram experiment to investigate the relationship between goal setting and unethical behavior. In our scenario studies we examine peoples' intuitions about goal setting and unethical behavior, and in our anagram experiment we examine the relationship between goals and actual behavior.

We recruited 889 participants to complete one of the eight scenario studies. In each study participants were asked to assess the relative likelihood that two similar individuals would engage in one of the following unethical actions: falsely report their billable hours, cheat on an exam, lie to a customer, and inflate an entry in a personal logbook. In each scenario participants compared two similar individuals with different types of goals.

Across the eight scenarios we manipulated whether or not individuals had a specific goal or no goal (an aim to "do your best"), whether the goal was met or unmet, how far the individual was from an unmet goal, and the incentive system related to the goal. With respect to goal-related incentives, we distinguish between <u>reward goals</u> that include a discrete incentive for reaching the goal, e.g. a bonus for selling 30 cars, and <u>mere goals</u> that lack a discrete incentive, e.g. a goal to finish writing a paper by Thursday (Heath, Larrick, & Wu, 1999)

Across the eight scenarios we find results consistent with our hypotheses. Our first hypothesis predicts that people with specific, unmet goals will be more likely to cheat than people without goals (people trying to "do their best"). In the first scenario, respondents expected a consultant

with a reward goal to be significantly more likely to lie about the number of hours he had actually worked than a similar consultant attempting to do his best, 79.4% versus 4.6% (16% thought both were equally likely), $\chi^2(1) = 87.3$, $\underline{p} < .001$. Similarly, in the second scenario a different group of respondents expected the consultant with a specific mere goal to be significantly more likely to lie about the number of hours he had actually worked than a similar consultant attempting to do his best, 56.7% versus 14.4% (28.8% thought both were equally likely), $\chi^2(1) = 26.2$, $\underline{p} < .001$. In the third scenario, most respondents expected a student with an unmet mere goal to be more likely to cheat on an exam than a similar student attempting to do her best, 79.6% versus 1.9% (18.5% thought both were equally likely), $\chi^2(1) = 80.2$, $\underline{p} < .001$.

Our second hypothesis predicts that people with specific, unmet goals will be more likely to cheat than people with specific, met goals. In the fourth scenario, the majority of respondents expected a salesperson with an unmet reward goal to be more likely to lie to a customer than a similar salesperson with a met reward goal, 78.8% versus 3.5% (17.7% thought both were equally likely), $\chi^2(1) = 77.7$, $\underline{p} < .001$. Similarly in the fifth scenario, a majority of respondents expected a salesperson with an unmet mere goal to be more likely to lie to a customer than a similar salesperson with a met mere goal, 84.6% versus 1.6% (13.8% thought both were equally likely), $\chi^2(1) = 98.2$, $\underline{p} < .001$. In the sixth scenario, respondents rated a jogger with an unmet goal as more than twice as likely to record an inflated number in his logbook than a similar jogger with a met goal, 58.2% versus 24.5% (17.3% thought both were equally likely), $\chi^2(1) = 15.0$, $\underline{p} < .001$.

Our third hypothesis considers the decision maker's proximity to the goal. We expect individuals who are very close to, but have not yet reached the goal to be more likely to lie or cheat than similar individuals who are distant from, but have not yet reached the goal. In scenarios seven and eight, a majority of respondents judged the salesperson closer to the goal as the person most likely to lie to a customer; this was true when the salespeople had unmet reward goals, 62% versus 2% (36% thought both were equally likely), $\chi^2(1) = 56.25$, $p < .001$, and when the salespeople had unmet mere goals, 60% versus 8% (32% thought both were equally likely), $\chi^2(1) = 39.76$, $p < .001$.

Results from these studies support our thesis linking goal setting to unethical behavior. Consistent with our first three hypotheses, we found that respondents (1) expected someone with an unmet goal to be more likely to engage in unethical behavior than someone attempting to do his or her best; (2) expected someone with an unmet goal to be more likely to engage in unethical behavior than someone with a met goal; and (3) expected someone with an unmet goal who is close to the goal to be more likely to engage in unethical behavior than someone with an unmet goal who is far from her goal. These hypotheses were supported for both mere and reward goals, suggesting that respondents believe that goals alone may motivate unethical behavior. Taken together, these results suggest that goal setting can predictably lead to unethical behavior, and this is likely to be true across a broad range of contexts, with or without financial incentives.

ANAGRAM EXPERIMENT

We extend our investigation to examine actual behavior in a laboratory study. We recruited 154 participants, and conducted a two-stage experiment. The first stage involved seven rounds of a

standard word creation task. In each round participants were instructed to create as many English words as they could from a list of seven letters within one minute. Participants were randomly assigned to one of three between-subject conditions: reward goal, mere goal, or do your best. Reward goal participants earned a financial reward ($2) for creating nine or more words per experimental round. Mere goal participants were given a goal of creating nine or more words per experimental round, but were offered no financial reward for attaining this goal. Do your best participants were told to do their best to create as many words as they could during each round. Both mere goal and do your best participants were given $10 at the beginning of the experiment for their participation, regardless of their performance. Reward goal participants were given an envelope with $14 at the beginning of the experiment, and asked to pay themselves and return any unearned money at the end of the experiment.

In the second stage of the experiment participants checked their own work and recorded the number of valid words they created on a separate answer sheet. In this stage of the experiment participants had an opportunity to misrepresent the number of words they had actually created. Once they had completed this task, participants placed their answer sheet and their workbook in separate sealed boxes. We used this method for submitting work to give participants the impression that we could not easily match their workbook with their answer sheets. Once the experiment was completed, we were able to match workbooks and answer sheets by matching a unique set of letters on the workbook and answer sheet. We were then able to compare the actual number of valid words each participant created with the number of words they claimed to have created.

For each experimental round we coded the congruence between the claims participants made and the number of valid words they actually created. We define under-statements as cases in which participants created 9 or more valid words, but claimed to have created fewer than 9 words. We define over-statements as cases in which participants created fewer than 9 valid words, but claimed to have created at least 9 words.

Both motivational factors and careless factors can explain over-statements. Only careless factors, however, can explain under-statements. We examine the proposition that goal setting increases over-statements by comparing over- and under-statement fractions across conditions. Specifically, we compare the fraction of times participants over-stated productivity relative to the number of times they missed the goal and the fraction of times participants under-stated productivity relative to the number of times they met the goal. If only careless factors contribute to over-statements, we would expect the fraction of over-statements to roughly equal the fraction of under-statements for each of the experimental conditions. In fact, this is what we find in the do your best condition. In the do your best condition the over-statement fraction was actually slightly less than the under-statement fraction (.026 vs. .035, $t(56) = -0.32$, $p = .316$). In the goal setting conditions, however, this is not the case. We find that the average over-statement fraction is significantly greater than the average under-statement fraction in the reward goal (.133 vs. .025, $t(52) = 2.65$, $p = 0.01$) and the mere goal conditions (.084 vs. 0.0, $t(43) = 2.56$, $p = 0.01$). For each participant we also subtracted their under-statement fraction from their over-statement fraction. Average difference scores were –0.009, 0.084, and 0.108 for the do your best, mere goal, and reward goal conditions. These values are significantly different ($F(2, 151) = 3.67$, $p < 0.05$).

In the next set of analyses we decompose these results by comparing the over-statement and under-statement fractions separately. If only careless factors influence the difference between actual and reported performance, then both under- and over-statements should be the same across the three experimental conditions. In the do your best, mere goal, and reward goal conditions the average over-statement fractions were 0.026, 0.084, and 0.133, respectively; these values are significantly different ($F(2,151) = 3.78$, $p < .05$) and in the predicted direction. We find no systematic differences across conditions with respect to under-statements. In the do your best, mere goal, and reward goal conditions the average under-statement fractions were 0.035, 0.0, and 0.025, respectively; these values are not significantly different ($F(2,151) = 0.793$, $p = .46$). Taken together, these results identify a significant motivational effect for participants in the goal conditions to over-state performance. We do not find this motivation effect for participants in the do your best condition.

We also examine over-stated behavior by comparing the number of participants in each condition who over-stated their productivity at least once. As before we find that participants with goals over-stated performance more than participants without goals. In the do your best, mere goal, and reward goal conditions 10.53%, 22.73%, and 30.19% of participants over-stated their productivity in at least one round, respectively. In paired comparisons, the difference between proportions for participants in the reward goal and do your best conditions was statistically significant ($\chi^2(1, N = 110) = 6.64$, $p < .01$); and the difference between proportions for participants in the mere goal and do your best conditions was marginally significant ($\chi^2(1, N = 101) = 2.77$, $p = .096$).

Finally, we examine the relationship between proximity to the goal and the likelihood of over-stating productivity. In this analysis we only considered rounds in which participants created fewer than 9 valid words, and hence had the opportunity to over-state their productivity. When participants created fewer than 9 valid words they either honestly missed the goal or over-stated their productivity. For each participant who over-stated their productivity at least once and honestly missed the goal at least once, we calculated the average number of valid words they had actually created in both types of rounds. This creates two scores for each individual that we use in paired t-tests. Supporting our hypothesis, participants created an average of 5.05 words in honest rounds and 7.36 words in over-stated rounds, $t(26) = 9.18$, $p < .001$. In fact, participants were most likely to over-state their productivity when they had created 8 valid words. Of the 59 cases in which participants over-stated their performance, 30 (50.85%) occurred when participants had created 8 valid words, and most of these occurred when participants were given a specific goal. Participants created 8 valid words in a total of 27, 29, and 33 rounds in the do your best, mere goal, and reward goal conditions, respectively. When we examine just these rounds, we find that participants claimed to have created 9 or more words 4 (14.81%), 10 (34.48%), and 16 (48.48%) times in the do your best, mere goal, and reward goal conditions, respectively.

Supporting our thesis, we find that participants who were given mere or reward goals over-stated their productivity significantly more often than participants who were asked to do their best. Interestingly, participants in this study with mere goals, who attained no monetary or social rewards for reaching a goal, were more likely to over-state their productivity than were

participants attempting to do their best. This suggests that goal setting alone, without economic incentives, increases the value people derive from over-stating productivity. In addition, we find that participants were most likely to over-state performance when they were close to reaching the goal.

GENERAL DISCUSSION

This article extends prior work by demonstrating that goals motivate not only constructive behavior, but also unethical behavior. We find that participants with goals were more likely to misrepresent their productivity than were participants attempting to do their best. This was true both for goals with and for goals without incentives. We also find that proximity to the goal influenced behavior. Participants were more likely to over-state performance when they were close to, rather than far from, reaching the goal.

Taken together, these results identify a serious "side-effect" to setting goals, and offer insight into the mechanics of this problem. Prescriptively, managers should be vigilant for unethical behavior and motivated communication (Schweitzer & Hsee, in press) when setting goals, particularly when employees are very close to reaching the goal. In these cases organizational controls and transparency may be particularly important. Finally, results from this work suggest that educators should include an ethics "warning" when they teach students about goal setting.

REFERENCES

Heath, C., Larrick, R. P., & Wu, G. 1999. Goals as reference points. **Cognitive Psychology,** 38: 79-109.

Ivancevich, J. 1974. Changes in performance in management by objectives program. **Administrative Science Quarterly**, 19: 563-574.

Kahneman, D., & Tversky, A. 1979. Prospect Theory: An analysis of decision under risk. **Econometrica**, 47: 263-91.

Locke, E. A., & Latham, G. P. 1990. **A Theory of Goal Setting and Task Performance**. Englewood Cliffs, NJ: Prentice Hall.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. 1981. Goal setting and task performance: 1969-1980. **Psychological Bulletin**, 90: 125-152.

Mento, A. J., Steel, R. P., & Karren, R. J. 1987. A meta-analytic study of the effects of goal setting on task performance: 1966-1984. **Organizational Behavior and Human Decision Processes**, 39: 52-83.

Schweitzer, M., & Hsee, C. (in press). Stretching the truth: Elastic justification and motivated communication of uncertain information, **Journal of Risk and Uncertainty**.