

Stretch Goals and the Distribution of Organizational Performance

Michael Shayne Gary (sgary@unsw.edu.au)
UNSW Business School, UNSW,
Sydney, NSW 2052 Australia

Miles M. Yang
Curtin Business School, Curtin University,
Bentley, Western Australia, Australia

Philip W. Yetton
Faculty of Business and Law
BL Deakin Business School

John D. Sterman
Sloan School of Management, MIT
Cambridge, MA 02142 USA

Abstract

Many academics, consultants and managers advocate stretch goals to attain superior organizational performance. However, existing theory speculates that, although stretch goals may benefit some organizations, they are not a ‘rule for riches’ for all organizations. To address this speculation, we use two experimental studies to explore the effects on the mean, median, variance, and skewness of performance of stretch compared with moderate goals. Participants were assigned moderate or stretch goals to manage a widely used business simulation. Compared with moderate goals, stretch goals improve performance for a few participants, but many abandon the stretch goals in favor of lower self-set goals, or adopt a survival goal when faced with the threat of bankruptcy. Consequently, stretch goals generate higher performance variance across organizations and a right-skewed performance distribution. Contrary to conventional wisdom, we find no positive stretch goal main effect on performance. Instead, stretch goals compared with moderate goals generate large attainment discrepancies that increase willingness to take risks, undermine goal commitment, and generate lower risk-adjusted performance. The results provide a richer theoretical and empirical appreciation of how stretch goals influence performance.

Key Words: Goals, aspirations, stretch objectives, performance variance, skewed distribution

Introduction

Many managers, consultants and academics advocate the use of stretch goals to boost organizational performance (Collins and Porras 1994; Kerr and Landauer 2004; Locke and Latham 2013; Thompson et al. 1997). These advocates have influenced the Boards of Directors and the top management teams of organizations to move away from goals based on routine adjustments to aspiration levels and, instead, to adopt explicit stretch goals for organizational performance (Fuller and Jensen 2010; Hamel and Prahalad 1993). Advocates argue that stretch goals improve performance by disrupting complacency, promoting new ways of thinking, stimulating search and innovation, energizing employees, and guiding effort and persistence (Shinkle 2011 provides a review).

These arguments are based on two assumptions that are subject to major potential validity threats. First, stretch goal advocates extend the generally accepted finding from the managerial psychology literature that challenging goals have a positive performance main effect on well structured tasks (for reviews of this research see Locke and Latham 1990; Locke and Latham 2013) to claim that stretch goals have a positive main effect on organizational performance. There is no empirical evidence or theoretical framework to support this generalization. Indeed, goal setting theory is described by Locke and Latham (1984) as “a motivational technique that works” but does not constitute a theory explaining how goals affect organizational performance.

Second, the cases illustrating the benefits of stretch goals (for example, Collins and Porras 1994; Kerr and Landauer 2004; Peters and Waterman 1982; Slater 1999; Thompson et al. 1997) are selected ex-post on the basis of success. Sampling on the dependent variable is a classic source of bias and a major internal validity threat. The cases do not constitute a random sample allowing reliable generalizations to the populations of organizations from which they were drawn.

Advocates of stretch goals disregard these two validity threats and claim that adopting stretch goals is a ‘rule for riches’ for organizations. Implicitly, if not explicitly, the assumption is that the distribution of organizational performance simply shifts to the right, improving performance, without significant effect on the variance or skewness of the distribution. If stretch goals only shifted the

distribution of expected performance to higher levels, few would argue against them and research would be needed only to identify the appropriate levels for stretch goals in different contexts.

However, by definition, stretch goals are difficult to achieve and are therefore likely to be achieved by only a few organizations (Sitkin et al. 2011). The many organizations that fall short will experience significant attainment discrepancies. Research shows that managers initially respond to attainment discrepancies by engaging in local search for incremental improvements of the current strategy (Martin and Mitchell 1998; Stuart and Podolny 1996) and/or in distant search for entirely new strategies through risky organizational changes (Greve 1998; Rosenkopf and Nerkar 2001).

In complex settings, new strategies often fail and can generate very low performance that threatens the organization's survival. Some managers then adopt low risk survival strategies (Boyle and Shapira 2012; March and Shapira 1987, 1992). Other managers adopt self-set aspiration levels below the exogenous stretch goals to resolve the attainment discrepancy. This dynamic adjustment process continues until the discrepancy is resolved when the self-set aspiration level equals actual performance (Lant 1992; Mezas et al. 2002).

These feedback processes suggest stretch goals, compared with more easily achievable goals, do not simply shift the performance distribution to the right. Instead, stretch goals alter the shape of the performance distribution. Therefore, in this paper we address the question: What are the effects of adopting stretch goals on the mean, variance, and skewness of organizational performance?

To address this question, we conduct two laboratory experiments employing a widely used, realistic business simulation. Participants take the role of senior management leading a start-up in a mature industry. We vary goals for financial performance to explore how stretch compared with moderate goals affect the performance distribution. In addition, to explain the drivers of these outcomes, we estimate the effects of stretch compared with moderate goals on goal commitment and willingness to take risk.

We find that stretch goals are not a rule for riches for all organizations. Instead, they lead to riches for a few. The cases used by the advocates to illustrate the success of stretch goals are ex-post members of 'the few.' The cases should not be used to justify the adoption of stretch goals ex-ante by 'the many,' for whom the adoption of stretch goals reduces risk-adjusted performance.

Research on Organizational goals

Goals or aspiration levels have long played an important role in organization theory (e.g., Cyert and March 1963; Simon 1964) and are central to understanding decision making in organizations (Boyle and Shapira 2012; Sitkin et al. 2011). Decision makers respond to attainment discrepancies in multiple ways, including local search for incremental improvement, distant search for novel strategies, aspiration adjustment, and the adoption of survival goals to limit risk when the organization's survival is threatened (Argote and Greve 2007; Lant and Shapira 2008; Lant 1992; Mezias et al. 2002).

We review the existing research to develop a causal theory that integrates these mechanisms and shows the interdependent feedback processes they create. Drawing on this analysis, we develop five hypotheses regarding the impact of stretch compared with moderate goals. The first three hypotheses specify the effects on expected performance, performance variance, and performance skewness. The other two hypotheses examine the effects of stretch goals on goal commitment and risk taking.

Local and Distant Search

Local search is defined as the search for solutions in the neighborhood of the current strategy, activities, or knowledge (Levinthal and March 1981; Rosenkopf and Nerkar 2001; Stuart and Podolny 1996). Decision makers assess the difference between goals and actual performance and are motivated to perform at or slightly above their goals (Lant and Shapira 2008). Local search emphasizes refinements to and greater efficiency in existing strategies, technologies, and activities.

Depending on the opportunities for improvement and success in implementation, local search can reduce the discrepancy between goals and performance (Levinthal and March 1981). For example, Martin and Mitchell (1998) show that local search motivated incumbents in the MRI diagnostic imaging equipment industry to introduce designs similar to those in existing products. Similarly, Stuart and Podolny (1996) show that patenting activity by semiconductor firms typically concentrated in the technological domains where the firm had previously patented.

In contrast, distant search is characterized by exploration of novel strategies rather than incremental improvements on the existing strategy and routines (Levinthal 1997). Rosenkopf and Nerkar (2001) show that many companies in the optical disk drive industry explored beyond their

current organization and/or technological boundaries. These distant searches had an important impact on subsequent technological evolution. In addition, Lant et al. (1992) show that poor performing organizations initiated strategic reorientations more frequently than did high performers, providing evidence that distant search increases with larger attainment discrepancies. Similarly, Greve (1998) shows that larger performance shortfalls relative to historical aspiration levels and peer performance increased the probability of strategic change in the radio broadcasting industry.

The advocates of stretch goals argue that superior, breakthrough performance is unlikely to result from local search or incremental improvement. Instead, they urge decision makers to set ambitious goals that cannot be achieved using variations on the existing strategies or routines:

A 'stretch target' is one that the organization cannot achieve simply by working a little harder or a little smarter. To achieve a stretch target, people have to invent new strategies, new incentives — entirely new ways of achieving their purpose. (Rose 2012)

The underlying rationale of stretch goals is that people have an innate desire to be challenged and that they will redouble their efforts when they are presented with a seemingly unattainable target. (Grinnell 2012)

So, to realize stretch goals, decision makers must try something radically new (Sitkin et al. 2011), continuing the search until a strategy is discovered that closes the attainment gap.

However, distant search, including strategic change, involves taking the risk that the new strategy will fail, leading to lower performance (Greve 1998). Indeed, empirical studies show high organizational risk taking is related to low subsequent performance relative to aspirations (Bowman 1982; Bromiley 1991; Bromiley et al. 2001). For example, Bromiley (1991) finds performance below the industry average leads to greater risk taking, which results in lower subsequent performance.

The ruggedness or smoothness of the performance landscape affects how hard it is to develop a successful new strategy from locally available information (Rivkin 2000). The greater the number of factors conditioning success and the more complex the interactions among them, the more rugged is the performance landscape. The more complex the payoff landscape, the higher the cost of distant search because it is harder to determine where to jump, reducing the probability of success (Gavetti & Levinthal 2002). In addition, delays between the selection of strategies and their impact, the presence or absence of nonlinearities, and the number and strength of feedback effects affect how difficult it is

to find better strategies (Rahmandad et al. 2009; Sterman 1989, 1994). Because decision makers do not know ex-ante the expected returns to search, they frequently over-search, lowering performance (Earley et al. 1989). Of course, distant search can be successful. The smoother the landscape and the greater the organization's skill in identifying promising strategies and implementing them successfully, the greater the probability that distant search boosts performance.

Aspiration Level Adaptation and the Survival Reference Point

Aspiration adaptation is another important dynamic process in the theory of organizational goals. Adapting aspiration levels reduces the attainment discrepancy as managers adopt self-set goals below the exogenous goals. Over time, this dynamic leads to convergence between current performance and the aspiration level (the goal adopted in practice). Instead of continuing to search and learn from performance feedback, decision makers judge current performance to be satisfactory.

Such goal erosion (Forrester 1968, 1975) is common. Lant (1992) finds that participants in a management simulation adjust their aspirations in response to performance feedback. In a field study of a large financial services company, Mezas et al. (2002) find strong effects on sales targets of the previous aspiration level and attainment discrepancy. Jordan and Audia (2012) show that, instead of continuing to search and learn from performance feedback, decision makers enhance their self-image by assessing current performance as satisfactory.

Decision makers' willingness to erode their goals is a negative function of goal commitment, defined as the determination to reach a specific goal (Hollenbeck and Klein 1987; Hollenbeck et al. 1989; Klein et al. 1999; Klein et al. 2001). The practitioner literature on stretch goals highlights the role of goal commitment (Collins and Porras 1994), arguing that aggressive goals reset an organization's aspirations, overcoming complacency and motivating both local search for improvements and distant search for novel strategies. However, repeated failure to achieve goals erodes commitment to exogenously imposed goals, resulting in downward aspiration adaptation towards current performance levels (Forrester 1968, 1975; Sterman 2000).

In the extreme case, when organizational survival is threatened, managers avoid taking additional risks and abandon distant search for a radical new strategy (March and Shapira 1987, 1992).

Abandoning distant search avoids the risk of choosing an ineffective strategy, decreasing the chance of organization failure. While focusing on survival by limiting search may avoid bankruptcy, it can also trap the organization near the survival point.

Stretch goals and organizational performance

As described above, theory and evidence show that the impact of goals on performance involves multiple, interacting nonlinear feedback processes, which we integrate in Figure 1.

[Insert Figure 1]

Consistent with the literature, the attainment discrepancy depends on the decision maker's aspiration level compared to actual performance. Importantly, the exogenous goal for performance and the endogenous aspiration level are distinct constructs. The exogenous goal for performance is set for decision makers by others, for example, a stretch goal set for managers by the CEO or Board. In contrast, the endogenous goal is *the goal adopted in practice* by the decision makers.

As described by Forrester (1968, 1975), the endogenous aspiration level is conditioned by both the exogenous goal and by past performance as aspirations adapt over time toward actual performance (Lant 1992). Following Forrester, the endogenous aspiration level can be modeled as a weighted average of (perceived) actual performance and the exogenous goal for performance, with goal commitment as the weight. When decision makers are fully committed to the exogenous goal, aspirations are unaffected by actual performance. As goal commitment falls, the aspiration level is increasingly contingent on past performance.

The attainment discrepancy creates the balancing “Local Search” feedback, B1, in Figure 1. Faced with an attainment discrepancy, managers undertake local search that, if successful, improves performance, thus reducing the attainment discrepancy.

The attainment discrepancy can also lead to distant search for more effective strategies. Consistent with the theory and evidence reviewed above, a large attainment discrepancy motivates greater willingness to take the risk of engaging in distant search (the positive link between the attainment discrepancy and willingness to take risk, and from willingness to take risk to distant search). Distant search can be successful or unsuccessful. If successful, the efficacy of the current

strategy improves, boosting performance and reducing the attainment discrepancy, forming the balancing “Distant Search” feedback, B2. However, when distant search leads to the selection of an ineffective strategy, the efficacy of the current strategy falls, increasing the attainment discrepancy and pressure to engage in further distant search. The result is the Strategy Churn reinforcing feedback (R1), in which low performance leads to still more distant search.

The probability of selecting an effective strategy determines which of these two feedbacks dominates. If the performance landscape is smooth and the organization’s search capabilities are strong, the higher the probability that loop B2 dominates and distant search leads to better strategies and performance improvement. Alternatively, the more rugged the landscape and the weaker the organization’s search capabilities, the higher the probability that loop R1 dominates and the selection of poor strategies leads to strategy churn.

At any point in time, decision makers must choose whether to continue with the current strategy, focus on local search for improvements (loop B1), or engage in distant search for a new strategy (loops B2 and R1). The relationships between the attainment discrepancy and each type of search are likely to be nonlinear because resources, including management attention, are limited. When the attainment discrepancy is small, managers are more likely to believe that local search will be sufficient to close the gap, reducing the need to engage in risky distant search. However, if managers believe the attainment gap is so large it cannot be closed through local search alone, more resources will be allocated to distant search.

The reinforcing “Aspiration Adjustment” feedback, R2, captures the adaptation of goals to recent performance as one pathway to close the attainment discrepancy. The strength of that loop is moderated by goal commitment, which is governed by the reinforcing “Ties to Exogenous Goal” feedback, R3. Large, persistent differences between actual performance and the exogenous goal erode commitment to the exogenous goal, which then shifts the aspiration level away from the exogenous goal towards actual performance, thus closing the attainment discrepancy.

Finally, the balancing “Survival Focus” feedback, B3, captures decision makers’ propensity to focus on survival when performance is very low. As performance approaches the survival point, the growing threat of bankruptcy reduces decision makers’ willingness to engage in distant strategy

search. This limits the risk of choosing an ineffective strategy, but also limits the opportunity to discover a better strategy.

Hypotheses regarding the distribution of performance under stretch goals

Although research on the effects of stretch goals on the distribution of organizational performance is limited, scholars have long sought to understand the variance and skewness of performance distributions (for example Denrell and Liu 2012; Levinthal 1991; March and Shapira 1992; March 1991; Simon 1955). Because of the complexity of the feedbacks in Figure 1, it is not obvious how stretch goals affect the performance distribution. However, consistent with some prior work, the presence of reinforcing feedbacks suggests stretch goals may increase the variance in performance by amplifying small initial differences in the attainment discrepancy and lead to a skewed performance distribution. For example, Denrell and Liu (2012) investigate the effects of skill, past success, and luck on the distribution of performance using both simulation and a laboratory experiment. They conclude, “When performance depends on past performance, errors and chance may not average out but may be amplified through rich-get-richer [i.e., reinforcing feedback] dynamics...that often generate fat-tailed distributions” (Denrell and Liu 2012: 9334).

Performance Variance. The feedback processes in Figure 1 have major implications for the effects of stretch goals on performance variance. When performance is safely above the survival reference point, performance below the aspiration level increases willingness to take risk and leads to more distant search for novel, untested strategies (Argote and Greve 2007; Bromiley et al. 2001; March and Shapira 1987). Stretch compared with moderate goals create larger attainment discrepancies and, therefore, higher willingness to take risk (Knight et al. 2001) through more extensive search for and trials of new strategies (Bromiley et al. 2001; Greve 1998, 2003). The distant search process (feedbacks B1 and R1) generates a wide range of potential strategies with different performance payoffs (Siggelkow and Rivkin 2006; Winter et al. 2007). Thus, stretch goals compared

with moderate goals should generate higher performance variance across organizations.¹

The reinforcing feedbacks – Strategy Churn R1, Aspiration Adjustment R2, and Ties to Exogenous Goal R3 – also increase performance variance by amplifying differences in initial conditions and among decision makers. The latter include differences in decision makers’ risk preferences, mental models and search heuristics, how they allocate their cognitive and other resources between local and distant search as the attainment discrepancy changes, and their propensity to erode their commitment to the stretch goals in the face of sustained differences between actual performance and the exogenous goal. Thus, the feedback structure in Figure 1 predicts:

Hypothesis 1: Stretch compared with moderate organizational goals generate higher variance in performance.

Skewness in Performance. Advocates of stretch goals typically assume high commitment to the formally assigned (exogenous) goal—that the stretch objective becomes the aspiration level guiding managerial behavior. The feedback processes in Fig. 1 suggest more complex dynamics. For managers who initially do well relative to the exogenous goal, performance improvement strengthens goal commitment, leading to stronger commitment to the goal and motivating additional effort—the reinforcing feedback R3 operates as a virtuous cycle. For managers who initially do poorly relative to the exogenous goal, however, the large discrepancy can cause goal commitment to erode, weakening local and distant search, causing the performance shortfall to persist, and further undermining goal commitment—feedback R3 operates as a vicious cycle.

Such goal erosion is particularly likely in complex environments where distant search frequently does not yield superior strategies (Gary and Wood 2011; Rivkin 2000); where it is difficult to determine which components of a strategy generate high performance (Fang 2012), and where a rugged strategy-performance landscape degrades the efficacy of search heuristics (e.g., Siggelkow and Rivkin 2006). In such environments stretch goals are likely to cause large and persistent performance shortfalls, eroding goal commitment and leading to downward aspiration adaptation to reduce the

¹ The focus here is on performance variance across a large sample of organizations at a particular time rather than performance variance over time for particular organizations; an important topic that has been examined in prior research (Bromiley 1991; Bromiley et al. 2001; Fiegenbaum and Thomas 1988).

attainment discrepancy (Lant 1992; Mezias et al. 2002). Over time, current performance levels become the aspiration levels despite the ambition of the stretch goals. Feedback loops R2, and R3 dominate, undermining the effectiveness of local and distant search (B1, B2).

In addition, for some decision makers, distant search results in the selection of ineffective strategies that reduce performance to the point of threatening the organizational survival. These decision makers shift to a survival goal (Boyle and Shapira 2012; March and Shapira 1987, 1992), converting a large attainment discrepancy (“I’m falling far short of the assigned goal”) into a small positive position (“I’m not dead yet”). Focused on survival (feedback B4), these decision makers minimize risk by abandoning distant search. Local search dominates (feedback B1), learning is limited, and mental models become rigid (Bromiley 1991; Bromiley et al. 2001; March and Shapira 1992). Although some still go bankrupt, others survive, trapped close to the survival reference point.

Consequently, even if managerial ability, risk attitudes, and other attributes are distributed normally in the population of decision makers, the feedback processes identified in Figure 1 amplify the success of those with the skill or luck to do well, stretching the performance distribution to the right, high-performance tail. The erosion of goal commitment and adoption of survival strategies truncates the left, low-performance tail of the distribution. These effects combine to generate right-skewed organizational performance distributions. Formally:

Hypothesis 2: Stretch compared with moderate organizational goals generate right-skewed performance distributions.

Expected Organizational performance. It is generally accepted in the goal setting literature that there is a positive main effect on performance of challenging goals compared with do your best goals (see reviews by Locke and Latham 1990, 2013). This assumption underpins recommendations for organizations to adopt stretch goals: advocates argue that stretch goals boost performance by disrupting complacency, promoting new ways of thinking, and increasing energy (Slater 1999; Thompson et al. 1997).

However, the goal setting literature also finds that the positive goal main effect decreases with task complexity (Earley et al. 1989; Wood et al. 1987). In a review, Ordóñez et al. (2009) also argue that the beneficial effects of stretch goals have been overstated, commenting that stretch goals can

inhibit learning, distort risk preferences, and reduce intrinsic motivation. Sitkin et al. (2011) argue that stretch goals reduce the performance of poor performing organizations with limited slack, which, paradoxically, are the organizations most likely to adopt stretch goals in an effort to improve.

Current theory is therefore inconclusive about the performance effect of stretch compared with moderate goals. If local and/or distant search are likely to be successful, then stretch goals should increase expected performance through the balancing feedback loops Local Search (B1) and Distant Search (B2). Alternatively, if low performance erodes goal commitment (Ties to Exogenous Goal, R3) and aspiration levels (Aspiration Adjustment, R2), survival threat inhibits distant search (Survival Focus, B3), or complexity leads to selection of unsuccessful strategies (Strategy Churn, R1), then stretch goals could reduce performance. As a default option, we adopt the generally accepted hypothesis of a positive performance effect of stretch compared with moderate organizational goals:

Hypothesis 3a: Stretch compared with moderate organizational goals increase performance.

If stretch goals increase the variance in outcomes through reinforcing feedbacks R1, R2 and R3 (Hypothesis 1), *risk-adjusted* performance could fall even if stretch goals compared with moderate goals increase mean performance. Risk-adjusted performance is frequently operationalized via the Sharpe ratio—the ratio of mean performance to its standard deviation (Sharpe 1994). Note that the Sharpe ratio is distinct from managerial perceptions about willingness to take risk in subsequent decisions. Even if stretch goals boost mean performance compared to moderate goals (i.e., if feedbacks B1 and B2 dominate), risk adjusted performance could rise or fall, depending on the change in the variance. Consistent with H3a, we adopt the hypothesis of a positive risk-adjusted performance main effect for stretch compared with moderate goals:

Hypothesis 3b: Stretch compared with moderate organizational goals increase risk-adjusted performance.

Figure 2 summarizes the hypothesized effects of stretch compared with moderate goals on the distribution of organizational performance. The top panel shows the conventional assumption regarding the impact of stretch goals on performance. Specifically, stretch goals shift the performance distribution to the right, while leaving the variance and skewness unchanged. The bottom half of Figure 2 shows the same hypothetical distribution for moderate goals and an alternative distribution

under stretch goals, with higher variance, right-skew, higher expected performance, and higher risk-adjusted performance resulting from the processes specified in Hypotheses 1, 2, 3a and 3b.

[Insert Figure 2]

Willingness to Take Risk and Goal Commitment. Next, we explore how stretch goals affect willingness to take risk and goal commitment. Stretch compared with moderate goals generate larger attainment discrepancies. Large and persistent attainment discrepancies affect willingness to take risk through feedbacks Distant Search (B2) and Strategy Churn (R1). Larger attainment discrepancies increase the willingness to take risk that comes with greater distant search for new strategies (Greve 1998; Lant and Shapira 2008) and increase risk taking more generally (Bromiley 1991; Bromiley et al. 2001; Larrick et al. 2009). Formally:

Hypothesis 4: Stretch compared with moderate organizational goals increase willingness to take risk.

As shown by the Ties to Exogenous Goal feedback R3 in Figure 1, large and persistent attainment discrepancies undermine goal commitment because decision makers come to believe that the goal is not attainable (Hollenbeck and Klein 1987). Goals have no motivational effect if there is no commitment to them (Klein et al. 2001). Lower goal commitment leads to greater willingness to erode goals until aspirations converge to actual performance (Lant 1992; Mezas et al. 2002). Formally:

Hypothesis 5: Stretch compared with moderate organizational goals decrease goal commitment.

Methodology

We test Hypotheses 1-5 through two experiments in which participants manage a simulated organization, with random assignment to stretch and moderate goal conditions. Study 1 tests Hypotheses 1, 2, 3a, and 3b. Study 2 replicates the tests for Hypotheses 1, 2, 3a, and 3b, and extends that analysis to test Hypotheses 4 and 5.

The Simulated Organization

The People Express simulation is an interactive, computer-based management simulation of an airline operating in a competitive market (Graham et al. 1992). Participants take the role of the senior

management of a start-up airline, making quarterly decisions for aircraft orders, hiring, average fare, marketing expenditure and service scope. The user interface provides participants multiple reports and graphs throughout the simulation to support decision making (see online appendix).

The simulation is based on a real organization, People Express Airlines, and replicates many aspects of the business decision-making environment faced by managers in the actual organization. The competitive context includes a large number of interdependent variables with multiple feedback effects, time delays, and nonlinear relationships (Graham et al. 1992). These features are characteristic of the complex environments managers face when making strategic decisions. The simulation has been utilized in previous research (Bakken et al. 1992; Graham et al. 1992), and has been widely used in MBA and executive teaching.²

Growing the simulated start-up airline requires coordinating fleet growth and hiring to maintain service quality. Without adequate staffing, growth erodes service quality, driving customers away and leading to financial losses. Building a skilled workforce is complicated by hiring and training delays, and through the impact of inexperienced employees on service quality, workload, burnout, and turnover. The coordination challenge often limits participant success as fleet growth causes service quality to fall. As load factor drops, growth must be halted to avoid bankruptcy (Graham et al. 1992).

Participants

In Study 1, 134 managers enrolled in an Executive MBA program participated in the simulation as a class exercise. They averaged 36 years of age and more than 10 years of work experience. They were randomly assigned to 50 teams with two or three members (34 teams of three and 16 teams of two). Teams were randomly assigned either a stretch profit goal ($n = 25$) or moderate profit goal ($n = 25$). There were no differences in performance between groups composed of two or three members. Ten teams had time for only two simulation rounds. Another was dropped from the analysis because the team made a data entry error. The analyses reported here are based on the 39 remaining teams.

² The simulation is available from: <http://www.strategydynamics.com/microworlds/people-express/>.

In Study 2, 59 students from a large university participated as individuals in the simulation. Participants averaged 22 years of age, and 53 percent were female. Most were undergraduates. Eighteen percent were economics majors, 24 percent management majors, with the others majoring in other fields. Participants were randomly assigned stretch ($n = 30$) or moderate ($n = 29$) profit goals.

Procedure

Cumulative net income is adopted as the organizational performance measure for each simulated firm. The stretch goal is based on the 90th percentile and the moderate goal is based on the 50th percentile performance levels achieved in pilot tests in which decision makers were instructed “Do Your Best to maximize Cumulative Net income.”

In Study 1, teams in the stretch (moderate) goal group were told, “The Board of Directors has set your Cumulative Net Income target equal to \$315 (\$60) million by the end of 10 years. This long term growth in profit will deliver the financial results that our shareholders expect.” In terms of ecological validity, explicit performance goals, including profit goals, are often set exogenously from the perspective of managers.

Participants were instructed to complete three simulation rounds of 40 quarters each (120 decision trials), comprising 30 years of simulated experience. Participants were not provided with any base rate performance data. After each trial, participants received outcome feedback on their results for that trial plus cumulative performance in all their trials to date. After each round of 40 quarters, the simulation was reset to the same initial values and the next round began.

In Study 2 there are four changes to the procedure followed in Study 1. First, rather than teams, each participant managed their simulated firm independently. Second, at the start of each lab session, the experimenter read the instructions aloud and participants then spent 25 minutes working through introductory exercises to familiarize themselves with the simulation.³ Each participant then received a memorandum outlining the performance goals for their simulated firm (see online appendix). After reviewing the memorandum, participants completed a questionnaire assessing their goal commitment

³ The exercises are available upon request.

and willingness to take risk prior to beginning the first simulation round. The same procedure was repeated for the second and third rounds. Third, in addition to the goal for Year 10 (the end of the simulation), participants were also assigned cumulative net income goals for each year, starting in Year 3. Assigning both short and long-term goals is more effective for complex tasks (Latham and Seijts 1999). Fourth, participants were told that they would be paid \$5 for participating in the experiment plus performance-based bonuses of \$2 for each intermediate year (3-9) in which their cumulative profit met or exceeded the goal for that year and \$6 if they met or exceeded the final goal. Participants meeting or exceeding their goals in all years of the three trials earned \$65 – a maximum of \$20 in each of the three rounds of the simulation, plus the \$5 participation payment.

All participants began with identical resources, and the simulation is deterministic. Therefore, any differences in performance arise entirely from differences in participants' decisions. The only difference is the goal condition assigned to each participant, isolating the effects of stretch compared with moderate goals.

Measures

Performance is measured by cumulative net income at the end of the third simulation round. For those experiencing bankruptcy (equity falling below zero), we use cumulative net income when bankruptcy occurs, which can be negative or positive.

Willingness to Take Risk and Goal Commitment were both assessed in Study 2 via online questionnaire (see online appendix). A six-item, task-specific measure of willingness to take risk in the next decision round was developed from well-established measures of perceived risk taking (Ganzach et al. 2008; Sitkin and Weingart 1995; Weber et al. 2002). Items included, for example, "How much risk will you take in your aircraft purchasing decisions?" A five-item measure of goal commitment was adapted from prior research (Klein et al. 2001). Items included, for example, "Quite frankly, I don't care if I achieve the annual goals or not." Participants were asked to complete the questionnaire before each simulation round (i.e., three times in total).

All hypotheses were tested using results for the third simulation round. A learning phase is frequently added to experimental studies when a complex task is not familiar to participants. The first

two learning rounds provide participants with an opportunity to get familiar with the user interface and the decisions they will make, and to build their understanding of the simulated competitive environment, which should help them achieve higher performance in the final round.

Results

Figure 3 shows the distribution of performance for Studies 1 (top panel) and 2 (bottom panel) in Year 10 for moderate and stretch goals. Visual inspection of the performance distributions for the stretch goal conditions in both studies shows higher variance than in the moderate goal conditions, and the stretch goal distributions also appear to be right-skewed.

[Insert Figure 3]

Hypothesis 1 is assessed with the Levene test for equality of variances between the stretch and moderate goal groups for Year 10 of the simulation. The Levene test does not require normality of the underlying data. Hypothesis 1, *Stretch compared with moderate organizational goals generate higher variance in performance*, is supported. In Study 1 those assigned stretch compared with moderate goals exhibit significantly higher performance variance (SD = \$219 million vs. \$113 million: $L[1, 37] = 8.29, p < .01$). In Study 2, participants assigned stretch goals exhibit significantly higher performance variance (SD = \$526.6 million versus SD = \$114.5 million; $L[1, 57] = 13.15, p < .01$).

To test Hypothesis 2, we use the robust Jarque-Bera test to assess the skewness of the performance distribution (Gelade et al. 2015). Hypothesis 2, *Stretch organizational goals generate right-skewed performance distributions*, is supported. In Study 1, the performance distribution for stretch goals is right-skewed (Jarque-Bera $T = 12.88, p = .000$). In contrast, the performance distribution for moderate goals in Study 1 is not skewed (Jarque-Bera $T = .18, p = .67$). Similarly, in Study 2 the performance distribution for the stretch goal condition is right-skewed (Jarque-Bera $T = 4.90, p = .03$), but is not skewed under moderate goals (Jarque-Bera $T = 0.00, p = 1.00$).

To test Hypothesis 3a we examine differences in both mean and median performance. In highly skewed distributions, median performance is a more appropriate measure of central tendency than mean performance because the median is less influenced by extreme outcomes. The Mann-Whitney nonparametric test does not require a normal distribution and, for continuous response variables, tests

for a statistically significant difference in medians. We also test for differences in mean performance between the two goal conditions with a t-test on the log of performance to reduce skew.

Hypothesis 3a, *Stretch compared with moderate organizational goals increase performance*, is not supported. In Study 1, neither median nor mean cumulative profit under stretch goals (\$6 million and \$144 million, respectively) are statistically significantly different from the moderate goal condition (mean \$120 million and median \$112 million): Mann-Whitney's $U = 180.00$, $z = -0.25$, $p = .81$ and $t = 0.73$, $p = 0.47$). The results are the same in Study 2, where median and mean cumulative profit under stretch goals (\$185 million and \$362 million, respectively) are not statistically significantly different from the moderate goal condition (\$138 million and \$149 million, respectively): Mann-Whitney's $U = 348.0$, $z = -1.32$, $p = 0.19$ and $t = -0.58$, $p = 0.56$). The results show no positive main effect of stretch goals on either median or mean performance.

To assess Hypothesis 3b, we calculate the Sharpe ratio: the ratio of mean performance to the standard deviation for all participants in each condition. The Sharpe ratio is widely used in empirical finance to assess risk-adjusted performance (Sharpe 1994). We evaluate the null hypothesis that the Sharpe ratios are equal in the two goal conditions by estimating the 95% confidence interval of the difference between the two ratios from 5,000 bootstrapping samples.

Hypothesis 3b, *Stretch compared with moderate organizational goals increase risk-adjusted performance*, is not supported. In Study 1, the Sharpe Ratio for the stretch goal condition is 0.66 compared with 1.00 for the moderate goal condition. The difference between the Sharpe Ratios is not significant ($\Delta = 0.334$, $p = 0.19$, two-tailed test). In Study 2, the Sharpe Ratio for the stretch goal group is 0.67 compared with 1.31 for the moderate goal group, a statistically significant difference ($\Delta = 0.595$, $p = 0.02$, two-tailed test). Contrary to Hypothesis 3b, stretch compared with moderate goals decrease risk-adjusted performance—the opposite of the direction hypothesized.

Hypothesis 4, *Stretch compared with moderate organizational goals increase willingness to take risk*, is supported. Figure 4 shows mean willingness to take risk for the moderate and stretch goal conditions prior to each of the three simulation rounds. Willingness to take risk for those in the stretch goal condition prior to the third simulation round is higher than for those in the moderate goal condition (stretch goals: $\mu = 6.24$; moderate goals: $\mu = 4.39$; $t(57) = -3.11$, $p < 0.01$).

[Insert Figure 4]

Hypothesis 5, *Stretch compared with moderate organizational goals decrease goal commitment*, is supported. Figure 5 shows mean goal commitment for the moderate and stretch goal conditions prior to each of the three simulation rounds. Prior to Round 1, goal commitment is the same for both groups (stretch goals: $\mu = 5.39$; moderate goals: $\mu = 5.75$; $t(57) = 1.22$, $p = 0.23$). Goal commitment declines during Rounds 1 and 2 for those assigned stretch goals. Prior to the third simulation round, goal commitment for decision makers assigned stretch compared with moderate goals is significantly lower (stretch goals: $\mu = 3.99$; moderate goals: $\mu = 6.12$; $t(57) = 4.18$, $p < 0.001$).

[Insert Figure 5]

We also tested the prediction that low performance relative to exogenous goals erodes goal commitment by estimating a regression model with the deviation from the exogenous goal as a predictor of goal commitment: $C_{r=3} = b_0 + b_1(G - P_{r=2}) + \epsilon$, where $C_{r=3}$ is commitment to the exogenous goal prior to round 3, G is the exogenous goal for cumulative profit at the end of the round, and $P_{r=2}$ is actual cumulative net income achieved at the end of round 2. The model is significant ($R^2 = 0.28$, $F = 23.59$, $p = .000$). As expected, the coefficient for the deviation from exogenous goal is negative and significant ($b_1 = -.003$, $p = .000$). Participants who perform poorly relative to the exogenously assigned goal decrease their commitment to that goal. A one standard deviation increase in the performance shortfall decreases goal commitment by 1.2 (on a scale from 0-10).

Additional analysis

To further investigate the causal model in Figure 1, we examine whether stretch compared with moderate goals lead to different levels of the attainment discrepancy, endogenous aspirations, local search, and distant search. We infer endogenous aspirations using our measure of goal commitment, and therefore examine these relationships using the data from Study 2.

The causal model (Figure 1) predicts stretch goals will lead to higher endogenous aspiration levels and higher attainment discrepancy compared to moderate goals. The attainment discrepancy, D , is the endogenous aspiration level, P^* , less actual performance, P : $D_r = P^*_r - P_r$. To infer endogenous aspiration levels, we use Forrester's (1968) model in which the aspiration level is a weighted average

of the exogenous goal and historical performance: $P_r^* = C_r G + (1 - C_r)H$, where C is commitment to the exogenous goal, G is the exogenous goal at the end of the round, and H is historical performance. Consistent with prior research (Forrester 1968), historical performance is modeled as an exponentially weighted moving average of past performance: $H_r = H_{r-1} + \alpha(P_{r-1} - H_{r-1})$, where α is the fractional adjustment rate.⁴

As expected, MANOVA shows the inferred endogenous aspiration levels between the two goal conditions are significantly different: the mean inferred aspiration levels are \$438 million for the stretch condition versus \$117 million for the moderate goal condition ($F = 99.68$, $p = .000$).

However, further investigation provides a more nuanced insight into the dynamics of goal adaptation. Moderate goals are easier to achieve, so goal commitment for these participants remains high, limiting goal erosion: the difference between the exogenous goal and mean imputed aspiration level is \$26 million (18% below the goal). In contrast, under stretch goals, the large initial shortfall in performance relative to the exogenous goal undermines goal commitment, leading to the adoption of self-set goals closer to actual performance: the difference between the exogenous goal and mean imputed aspiration level is \$382 million (47% below the goal). While those in the stretch goal condition eroded their endogenous aspiration levels far more than those in the moderate goal condition, the aspiration levels for those assigned stretch compared with moderate goals remain significantly higher prior to round 3. Also, MANOVA shows that stretch compared with moderate goals result in significantly higher attainment discrepancy ($F = 74.68$, $p = .000$) prior to round 3. This is consistent with the causal model. Higher endogenous aspiration levels lead to higher attainment discrepancy for those assigned stretch goals.

Turning to local and distant search, we operationalized distant search as the number of decisions in a given quarter that changed by more than 30% compared with the previous quarter. Changes in any of the five decisions (aircraft orders, hiring, fares, marketing expenditure and service scope) are counted. Local search is operationalized as the average number of decisions in a given quarter

⁴ Since H is not observed, the adjustment rate α cannot be estimated. We assume $\alpha = 1$, which means $H_r = P_{r-1}$. The results are robust to other values.

changed more than zero but no more than 30%. To compare those who completed the full round (40 quarters) with those who experienced bankruptcy, we average the totals over the number of quarters completed. The causal model implies that stretch goals, by creating larger attainment discrepancies, motivate greater willingness to take risks, increasing distant search. Indeed, MANOVA shows stretch goals lead to significantly higher levels of distant search ($F = 12.23$, $p = 0.001$). Also consistent with the causal model, MANOVA shows no significant differences in local search between the two goal conditions ($F = .01$, $p = .91$): although rising attainment discrepancies motivate local search, the very large attainment discrepancies induced by stretch goals lead to a shift from local to distant search.

Discussion

This paper makes three contributions to research on organizational goals. First, the paper provides a theoretical framework integrating different concepts and empirical relationships that moderate the interactions between goals and performance. These include the motivational effects of stretch goals (Locke and Latham 2013), local and distant strategy search (Rosenkopf and Nerkar 2001), goal commitment (Klein et al. 2001), managerial intentions to take risk (Ganzach et al. 2008; Sitkin and Weingart 1995), performance variation (Bromiley et al. 2001; Hu et al. 2011), aspiration level adjustments (Lant 1992; Mezas et al. 2002), and the effect of the survival reference point (Boyle and Shapira 2012; March and Shapira 1987, 1992). Our efforts to develop an integrated theoretical framework respond to calls by scholars for research that assesses how these constructs interact and empirically tests the conventional wisdom (Hu et al. 2011). The causal theory in Fig. 1 shows how these multiple, interdependent feedback effects influence performance over time.

Second, studies 1 and 2 show that, contrary to conventional wisdom, stretch compared with moderate goals do not increase performance for the average or median organization (Hypothesis 3a is rejected) and *decrease* risk-adjusted performance (the opposite of Hypothesis 3b). Also, compared with moderate goals, stretch goals increase the variance in organizational performance (Hypothesis 1) and the skewness of organizational performance (Hypothesis 2).

We find no positive main effect of stretch goals on performance because only a few performers meet or exceed the stretch goals – 19 percent in Study 1 and 13 percent in Study 2 – with the rest

ending up either far below the stretch goals or near bankruptcy. Participants in the middle, who are far below the stretch goals, eroded their endogenous aspiration levels until they converged with actual performance (Lant 1992; Mezas et al. 2002). At that point, there was no attainment discrepancy and performance stabilized. The group of low performers who go bankrupt or achieve performance just above bankruptcy sought to reduce risk taking by limiting search to local, incremental changes (Boyle and Shapira 2012; March and Shapira 1987, 1992). Most of these low performers remain trapped in the low performance region; some go bankrupt despite abandoning distant search.

Our results are consistent with the simulations reported by Hu et al. (2011) connecting ambitious (i.e., stretch goal) strategies and ex ante risk taking. Similar to the Hu et al. simulations, our experiments show that that stretch goals initially generate large attainment discrepancies that motivate distant search. However, in the long run, many of these decision makers fail, or end up near the survival point and limit subsequent risk taking. This traps many of them in the low performance region. Our findings extend the Hu et al. (2011) results by highlighting another long run response to stretch goals: downward aspiration adjustment towards actual performance. Goal erosion was not included in the Hu et al. simulation model because the focal firm's historical performance was replaced by the performance of a reference group (either the top 10% of firms, the industry average, or the group performing just above the focal firm).

Advocates of stretch goals also implicitly assume that the variance in and skewness of the outcome distribution remain constant. Contrary to the assumption of constant performance variance in goal setting theory (see review by Locke and Latham 1990), our results show this assumption is invalid. The findings support speculations by Mosakowski (1998) and Sitkin et al. (2011) that stretch goals benefit some organizations but not others. Variance in performance is an endogenous outcome of complex feedbacks among explicit goals, goal commitment, local and distant search, survival threat and other factors. It should not be treated as errors distributed around a goal main effect.

Of course, higher performance variance could be attractive if the positive performance main effect is large enough. Risk-seeking investors prefer more volatility in returns in exchange for higher average return. However, the results show that stretch goals do not lead to higher expected performance or higher risk-adjusted performance. Instead, Study 2 shows stretch goals generate a

significant reduction in risk-adjusted performance. Risk-adjusted performance in Study 1 was also lower for stretch compared with moderate goals, although the difference in performance is non-significant.⁵

These results test and extend Mosakowski's (1998) theoretical analysis that questions whether stretch goals are a rule for riches for organizations. In both Studies 1 and 2, stretch goals generate riches for the few. The few high performers in the stretch goal condition identify and successfully execute profitable growth strategies. These strategies are hard to discover in dynamic environments characterized by time delays, feedback effects, nonlinearities, and combinatorial complexity (Gary and Wood 2011; Rivkin 2000). Only the top 19 percent in Study 1 and 13 percent in Study 2 achieve or nearly achieve the stretch goals. In contrast, for the vast majority, stretch goals frequently lead to strategy churn, low performance and erosion of goal commitment, and, in the extreme case, to survival threats that undermine the search for new strategies.

The third contribution extends prior theory by proposing and testing specific mechanisms that lead to the observed empirical results. The causal diagram in Figure 1 shows multiple interacting feedback loops that explain how stretch goals affect the first three moments of the performance distribution. Although Studies 1 and 2 were not designed to test each relationship in Figure 1, the results provide support for many of the proposed causal mechanisms.

The individual mechanisms in Figure 1 are not new in the literature. However, they have not been integrated into a single theoretical framework and many have not been tested empirically. We synthesized the mechanisms examined in prior work, showing how they interact to create multiple reinforcing and balancing feedback processes that interact nonlinearly with each other and with the complexity of the task environment to influence the effects of stretch goals on performance.

The interactions among these feedback processes in Figure 1 explain how stretch goals generate a right-skewed performance distribution. This is a novel finding in the literature on organizational

⁵ The stronger results in Study 2 may be at least partially due to differences in risk-taking behavior of individuals versus groups; groups are typically more risk averse than individuals (Maslet et al. 2009).

goals. It builds on and generalizes research that shows rich-get-richer reinforcing feedback processes lead to right-skewed performance distributions (Denrell and Liu 2012).

Figure 1 also shows how shifting to a survival goal curtails risk taking and increases the likelihood of becoming trapped near the survival point (Boyle and Shapira 2012). In this way, the survival mechanism plays an important role in determining the shape of the performance distribution. If decision makers did not care about organization survival, the large and persistent attainment discrepancies created by stretch goals would lead to greater risk taking and more bankruptcies. Performance variance would still increase but the distribution would be more symmetric.

Implications for Practice

Our findings have three implications for practice. First, motivated by a few highly successful cases, Boards and CEOs of publicly listed companies increasingly adopt stretch goals for financial performance. Two quotes illustrate the sentiment.

"Organization experience has demonstrated, the intelligent use of difficult or stretch goals can dramatically improve productivity, efficiency, and profitability" (Kerr and Landauer 2004: p. 134).

"It [setting stretch goals] doesn't mean that we know how we are going to get there, but at least we've got every human factor lined up and trying to achieve the targeted goal" (Denning 2012).

The results of both studies 1 and 2 highlight the attraction of stretch goals. The performance of the few who adopt stretch goals *and* are successful are likely to be very salient compared with the performance of others. For example, the top 20 percent of performers in the stretch goal condition in Study 2 achieved average cumulative net income ($\mu = \$952.85$ million) nearly a factor of eight greater than the average performance of the others in the stretch goal condition ($\mu = \$120.37$ million). The argument that successful cases are evidence for the efficacy of stretch goals is subject to the major validity threat of sampling on success ex-post, and then generalizing from a small, non-random sample to the population of organizations adopting stretch goals. This would be equivalent to selecting the top 20 percent of participants in Studies 1 and 2 ex post, while ignoring the rest, to demonstrate the benefits of stretch goals.

Second, instead of being evidence that organizations should adopt stretch goals, the small number of successful cases held up as exemplars for the benefits of stretch goals is evidence that stretch goals

are not a rule for riches for all organizations. Instead, they may enrich a few, while many more do not benefit or suffer. Without knowledge of the impact of different goal levels on the distribution of performance, it is not possible for managers to make informed decisions about the appropriate level of the goals that they should set for their organizations. The problem is compounded by the fact that people are unrealistically optimistic about their position in a distribution of peers on almost any positive trait or ability (MacCrimmon and Wehrung 1986).

Third, the findings inform the issue of setting appropriate goals for specific contexts. In particular, the results show that whether boards or top management should impose stretch goals on their organization depends on their attitudes towards risk. Those with large appetites for risk may still prefer stretch goals. However, for those who are risk-neutral or risk averse, stretch goals may not be desirable because the increase in performance variance – including the risk of failure – and the lower risk-adjusted return achieved by the typical organization outweigh the chances for improvement achieved by a few successful high performers.

In some situations stretch goals that lead to only a small number of highly successful organizations may be desirable. In venture capital or private equity, the value created by big winners, for example Apple and Amazon.com, can more than offset the poor returns or losses on the majority of organizations in the portfolio. In other settings, the higher risks associated with stretch goals are not appropriate. For example, stretch goals may not be appropriate in a medium-sized family-owned business that constitutes the majority of the family's net worth.

Preference for stretch versus moderate goals may also be contingent on the nature of the market. In markets characterized by reinforcing feedbacks, such as increasing returns, that lead to winner-take-all dynamics (Sternman 2000, Ch. 10), stretch goals may prove the only path to success: firms must “go for broke or die trying.” However, in markets where multiple firms can coexist, the risks of failure due to stretch goals may dominate, and the watchword should be “live and let live.”

These arguments show how the appropriate aggressiveness of goals depends on the context. In favorable markets, for example those with less rugged performance landscapes or high munificence, search for high performing strategies could be successful and bankruptcy unlikely (Denrell and March

2001). In contrast, distant search for high performing strategies in markets characterized by rugged performance landscapes is more likely to result in lower performance for most organizations.

Stretch goals may also shape the behavior of the individuals in organizations that then affect organizational outcomes. For example, increased variance and thus a more visible upper tail of high performers could serve as exemplars of success, motivating others in the organization to strive for higher performance, or lead to jealousy, the threat of being perceived as low-performing compared to a few stars, or other emotions that disrupt collaboration and the pursuit of organizational goals in favor of individual goals. These internal organizational effects would also depend on the structure of the internal organizational market. If one big hit is much better internally (e.g., through financial rewards and career advancement) than five moderate outcomes, then managers may want to go for stretch goals. In situations where extreme effort and risk are subject to decreasing returns or the consequences of failure to meet targets is large, then stretch goals would be problematic.

Which of these possible individual impacts dominates depends on who carries the increased risk created by stretch goals. Boards and shareholders are unlikely to accept repeated failure by the CEO to achieve stretch goals. So, with 80% or more failing to achieve the stretch goals, do stretch goals increase CEO turnover? Similarly, if the CEO of a multi-business organization sets stretch goals for the division heads, how does the CEO hold them accountable, without becoming subject to the high cost of executive churn?

Limitations and Extensions

Like all studies, our study has limitations. First, we tested only two goal levels: stretch and moderate goals. However, organizations adopt goals that span the continuum from easy, through stretch, to, perhaps, impossibly difficult goals. We also held market conditions and context constant. Future research can explore more gradations in goals and different market and industry contexts, including variations in the complexity of the environment.

Second, we manipulated only one type of goal: cumulative profit. In practice, managers face multiple goals, such as profit, growth, share price and market share (Short and Palmer 2003). Multiple goals require tradeoffs among resources including managerial attention (Ethiraj and Levinthal 2009).

It seems unlikely the additional complexity of multiple goals compared to a single goal would generate a positive stretch goal main effect, but research should confirm this. Also, the effects of stretch goals on different outcome measures, such as idea generation, should be examined.

Third, the goals participants faced in our experiments were fixed. Participants who fell behind early in a round faced a more difficult situation than those who did well early. In many organizations, goals are adaptive (Mezias et al. 2002). If boards or CEOs impose stretch goals that prove so difficult that they undermine goal commitment, they, or their successors, may choose to lower those goals, hoping that more realistic goals would improve performance by preventing large attainment discrepancies from eroding goal commitment and morale or lead to the adoption of risky new strategies or other actions that harm performance. Alternatively, however, adaptive high-level goals could lead to underperformance through goal erosion (Sterman 2000, Ch. 15).

Fourth, each participant worked independently and did not receive feedback on the performance of others. In reality, firms frequently compare themselves to the performance of similar firms in their industry. Including social comparisons in aspirations adjustments would be a useful extension of our theoretical framework and could be tested experimentally in future research.

Fifth, although Hypotheses 4 and 5 are supported and the additional analysis reported for Study 2 is consistent with the causal theory we offer to explain the effects of stretch goals, further research is needed to explore these feedbacks, including nonlinearities that affect which feedbacks dominate the dynamics of the system and how they are conditioned by individual and contextual factors.

Finally, stretch goals may lead to other unintended, harmful effects, including capability traps (Repenning and Sterman 2002), corrosion of organization culture, unethical behavior and illegal activity (Ordóñez et al. 2009), and burnout, mental illness and even suicide. Examples abound, and we urge scholars to undertake research to explore them.

Notwithstanding these opportunities for additional research, the results show that the effects of stretch goals are more complex than previous research indicates, and subject to more caveats and nuances than many practitioner advocates of stretch goals acknowledge. At the very least, decision makers should not simply assume that stretch goals significantly increase performance, but may also expose organizations, and those who work in them, to unintended, negative consequences.

References

- Argote, L., H.R. Greve. 2007. A behavioral theory of the firm—40 years and counting: Introduction and impact. *Organization Science* **18**(3) 337-349.
- Bakken, B., J. Gould-Kreutzer, D. Kim. 1992. Management flight simulators and organizational learning: some experimental evidence. *European Journal of Operational Research* **59** 167-182.
- Bowman, E.H. 1982. Risk seeking by troubled firms. *Sloan management review* **23**(4) 33.
- Boyle, E., Z. Shapira. 2012. The Liability of Leading: Battling Aspiration and Survival Goals in the Jeopardy! Tournament of Champions. *Organization Science* **23**(4) 1100-1113.
- Bromiley, P. 1991. Testing a causal model of corporate risk taking and performance. *Academy of Management Journal* **34**(1) 37-59.
- Bromiley, P., K. Miller, D. Rau. 2001. Risk in strategic management research. M. Hitt, R. Freeman, J. Harrison, eds. *The Blackwell Handbook of Strategic Management*. Blackwell Publishers, Malden, MA, 259-288.
- Collins, J., J. Porras. 1994. *Built to last: Successful habits of visionary companies*, 3rd ed. Random House Business, London UK.
- Cyert, R., J. March. 1963. *A Behavioral Theory of the Firm*. Blackwell Publishers Inc., Cambridge, MA.
- Denning, S. 2012. *In Praise Of Stretch Goals*. Forbes Retrieved October 30, 2014, from <http://www.forbes.com/sites/stevedenning/2012/04/23/in-praise-of-stretch-goals/>
- Denrell, J., C. Liu. 2012. Top performers are not the most impressive when extreme performance indicates unreliability. *Proceedings of the National Academy of Sciences* **109**(24) 9331-9336.
- Denrell, J., J. March. 2001. Adaptation as information restriction: The hot stove effect. *Organization Science* **12**(5) 523-538.
- Earley, P.C., T. Connolly, G. Ekegren. 1989. Goals, strategy development, and task performance: Some limits on the efficacy of goal setting. *Journal of Applied Psychology* **74**(1) 24-33.
- Ethiraj, S., D. Levinthal. 2009. Hoping for A to Z while rewarding only A: Complex organizations and multiple goals. *Organization Science* **20**(1) 4-21.
- Fang, C. 2012. Organizational Learning as Credit Assignment: A Model and Two Experiments. *Organization Science* **23**(6) 1717-1732.
- Fiegenbaum, A., H. Thomas. 1988. Attitudes toward risk and the risk-return paradox: Prospect theory explanations. *Academy of Management Journal* **31**(1) 85-106.
- Forrester, J.W. 1968. Market Growth as Influenced by Capital Investment. *Industrial Management Review (MIT) (currently published as the Sloan Management Review)* **9**(2) 83-105.
- Forrester, J.W. 1975. Planning and Goal Creation Ch. 10 *In Collected Papers of Jay W. Forrester*. Wright Allen Press, Cambridge, MA, 167-174.
- Fuller, J., M. Jensen. 2010. Just say no to Wall Street: Putting a stop to the earnings game. *Journal of Applied Corporate Finance* **22**(1) 59-63.
- Ganzach, Y., S. Ellis, A. Pazy, T. Ricci-Siag. 2008. On the perception and operationalization of risk perception. *Judgment and Decision Making* **3**(4) 317-324.
- Gary, M.S., R.E. Wood. 2011. Mental models, decision rules, and performance heterogeneity. *Strategic Management Journal* **32**(6) 569-594
- Gelade, W., V. Verardi, C. Vermandele. 2015. Time-efficient algorithms for robust estimators of location, scale, symmetry, and tail heaviness. *Stata Journal* **15**(1) 77-94.
- Graham, A., J. Morecroft, P. Senge, J. Serman. 1992. Model-supported case studies for management education. *European Journal of Operational Research* **59**(1) 151-166.
- Greve, H.R. 1998. Performance, Aspirations, and Risky Organizational Change. *Administrative Science Quarterly* **43** 58-86.
- Greve, H.R. 2003. Investment and the behavioral theory of the firm: Evidence from shipbuilding. *Industrial and Corporate Change* **12**(5) 1051-1076.
- Grinnell, J. 2012. *Setting Stretch Goals... and Avoiding SNAP Goals*. PM Hut Retrieved October 30, 2014, from <http://www.pmhut.com/setting-stretch-goals-and-avoiding-snap-goals>
- Hamel, G., C.K. Prahalad. 1993. Strategy as stretch and leverage. *Harvard Business Review* **71** 75-84.

- Hollenbeck, J.R., H.J. Klein. 1987. Goal commitment and the goal-setting process: Problems, prospects, and proposals for future research. *Journal of Applied Psychology* **72**(2) 212.
- Hollenbeck, J.R., H.J. Klein, A.M. O'Leary, P.M. Wright. 1989. Investigation of the construct validity of a self-report measure of goal commitment. *Journal of Applied Psychology* **74**(6) 951-956.
- Hu, S., D. Blettner, R.A. Bettis. 2011. Adaptive aspirations: performance consequences of risk preferences at extremes and alternative reference groups. *Strategic Management Journal* **32**(13) 1426-1436.
- Jordan, A., P. Audia. 2012. Self-enhancement and learning from performance feedback. *Academy of Management Review* **37**(2).
- Kerr, S., S. Landauer. 2004. Using stretch goals to promote organizational effectiveness and personal growth: General Electric and Goldman Sachs. *Academy of Management Executive* **18**(4) 134-138.
- Klein, H.J., M.J. Wesson, J.R. Hollenbeck, B.J. Alge. 1999. Goal commitment and the goal-setting process: conceptual clarification and empirical synthesis. *Journal of Applied Psychology* **84**(6) 885.
- Klein, H.J., M.J. Wesson, J.R. Hollenbeck, P.M. Wright, R.P. DeShon. 2001. The Assessment of Goal Commitment: A Measurement Model Meta-Analysis. *Organizational Behavior and Human Decision Processes* **85**(1) 32-55.
- Knight, D., C.C. Durham, E.A. Locke. 2001. The Relationship of Team Goals, Incentives, and Efficacy to Strategic Risk, Tactical Implementation, and Performance. *Academy of Management Journal* **44**(2) 326-338.
- Lant, T., Z. Shapira. 2008. Managerial reasoning about aspirations and expectations. *Journal of Economic Behavior & Organization* **66**(1) 60-73.
- Lant, T.K. 1992. Aspiration level adaptation: an empirical exploration. *Management Science* **38** 623-644.
- Lant, T.K., F.J. Milliken, B. Batra. 1992. The role of managerial learning and interpretation in strategic persistence and reorientation: An empirical exploration. *Strategic Management Journal* **13**(8) 585-608.
- Larrick, R., C. Heath, G. Wu. 2009. Goal-induced risk taking in negotiation and decision making. *Social Cognition* **27**(3) 342-364.
- Latham, G., G. Seijts. 1999. The effects of proximal and distal goals on performance on a moderately complex task. *Journal of Organizational Behavior* **20**(4) 421-429.
- Levinthal, D.A. 1991. Random Walks and Organizational Mortality. *Administrative Science Quarterly* **36**(3).
- Levinthal, D.A. 1997. Adaptation on Rugged Landscapes. *Management Science* **43** 934-950.
- Levinthal, D.A., J.G. March. 1981. A Model of Adaptive Organizational Search. *Journal of Economic Behavior and Organization* **2**(4) 307-333.
- Locke, E.A., G.P. Latham. 1984. *Goal setting: A motivational technique that works!* Prentice Hall, Englewood Cliffs, NJ.
- Locke, E.A., G.P. Latham. 1990. *A theory of goal setting and task performance*. Prentice-Hall, Englewood Cliffs, NJ.
- Locke, E.A., G.P. Latham. 2013. *New developments in goal setting and task performance*. Routledge Academic.
- MacCrimmon, K., D. Wehrung. 1986. *Taking Risks*. Free Press, New York.
- March, J., Z. Shapira. 1987. Managerial Perspectives on Risk and Risk Taking. *Management Science* **33**(11) 1404-1418.
- March, J., Z. Shapira. 1992. Variable risk preferences and the focus of attention. *Psychological Review* **99**(1) 172-183.
- March, J.G. 1991. Exploration and exploitation in organizational learning. *Organization Science* 71-87.
- Martin, X., W. Mitchell. 1998. The influence of local search and performance heuristics on new design introduction in a new product market. *Research Policy* **26**(7) 753-771.
- Maslet, D., N. Colombier, L. Denant-Boemont, Y. Lohéac. 2009. Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers. *Journal of Economic Behavior & Organization* **70**(3) 470-484.

- Mezias, S.J., Y.R. Chen, P.R. Murphy. 2002. Aspiration-level adaptation in an American financial services organization: A field study. *Management Science* 1285-1300.
- Mosakowski, E. 1998. Managerial prescriptions under the resource-based view of strategy: The example of motivational techniques. *Strategic Management Journal* **19**(12) 1169-1182.
- Ordóñez, L.D., M.E. Schweitzer, A.D. Galinsky, M.H. Bazerman. 2009. Goals gone wild: The systematic side effects of overprescribing goal setting. *Academy of Management Perspectives* **23**(1) 6-16.
- Peters, T., R. Waterman. 1982. *In Search of Excellence: Lessons from America's Best-Run Corporations*. Harper and Row, New York.
- Rahmandad, H., N. Repenning, J.D. Sterman. 2009. Effects of Feedback Delay on Learning. *System Dynamics Review* **25**(4) 309-338.
- Repenning, N., J. Sterman. 2002. Capability traps and self-confirming attribution errors in the dynamics of process improvement. *Administrative Science Quarterly* **46**(2) 265-295.
- Rivkin, J.W. 2000. Imitation of Complex Strategies. *Management Science* **46** 824-844.
- Rose, F. 2012. *Stretch goals require new way of doing business*. Federal News Radio Retrieved October 30, 2014, from <http://www.federalnewsradio.com/538/2693221/Stretch-goals-require-new-way-of-doing-business>
- Rosenkopf, L., A. Nerkar. 2001. Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry. *Strategic Management Journal* **22**(4) 287-306.
- Sharpe, W.F. 1994. The sharpe ratio. *Journal of portfolio management* **21** 49-49.
- Shinkle, G.A. 2011. Organizational Aspirations, Reference Points, and Goals: Building on the Past and Aiming for the Future. *Journal of Management* 415-455.
- Short, J., T. Palmer. 2003. Organizational performance referents: An empirical examination of their content and influences. *Organizational Behavior and Human Decision Processes* **90**(2) 209-224.
- Siggelkow, N., J.W. Rivkin. 2006. When exploration backfires: Unintended consequences of multilevel organizational search. *Academy of Management Journal* **49**(4) 779-795.
- Simon, H.A. 1955. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics* **69** 99-118.
- Simon, H.A. 1964. On the concept of organizational goal. *Administrative Science Quarterly* **9**(1) 1-22.
- Sitkin, S.B., K.E. See, C.C. Miller, M. Lawless, A. Carton. 2011. The Paradox of Stretch Goals: Organizations in Pursuit of the Seemingly Impossible. *Academy of Management Review* **36**(3) 544-566.
- Sitkin, S.B., L.R. Weingart. 1995. Determinants of risky decision-making behavior: A test of the mediating role of risk perceptions and propensity. *Academy of Management Journal* **38**(6) 1573-1592.
- Slater, R. 1999. *Jack Welch and the GE way: Management insights and leadership secrets of the legendary CEO*. McGraw-Hill Companies, New York.
- Sterman, J.D. 1989. Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science* **35**(3) 321-339.
- Sterman, J.D. 1994. Learning in and about complex systems. *System Dynamics Review* **10**(2-3) 291-330.
- Sterman, J.D. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, New York.
- Stuart, T.E., J.M. Podolny. 1996. Local search and the evolution of technological capabilities. *Strategic Management Journal* **17**(S1) 21-38.
- Thompson, K.R., W.A. Hochwarter, N.J. Mathys. 1997. Stretch targets: What makes them effective? *Academy of Management Executive* **11**(3) 48-60.
- Weber, E.U., A.R. Blais, N.E. Betz. 2002. A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making* **15**(4) 263-290.
- Winter, S.G., G. Cattani, A. Dorsch. 2007. The value of moderate obsession: Insights from a new model of organizational search. *Organization Science* **18**(3) 403-419.
- Wood, R., A. Mento, E. Locke. 1987. Task complexity as a moderator of goal effects: A meta-analysis. *Journal of Applied Psychology* **72**(3) 416-425.

Figure 1 Causal diagram showing feedbacks involved in organizational responses to goals

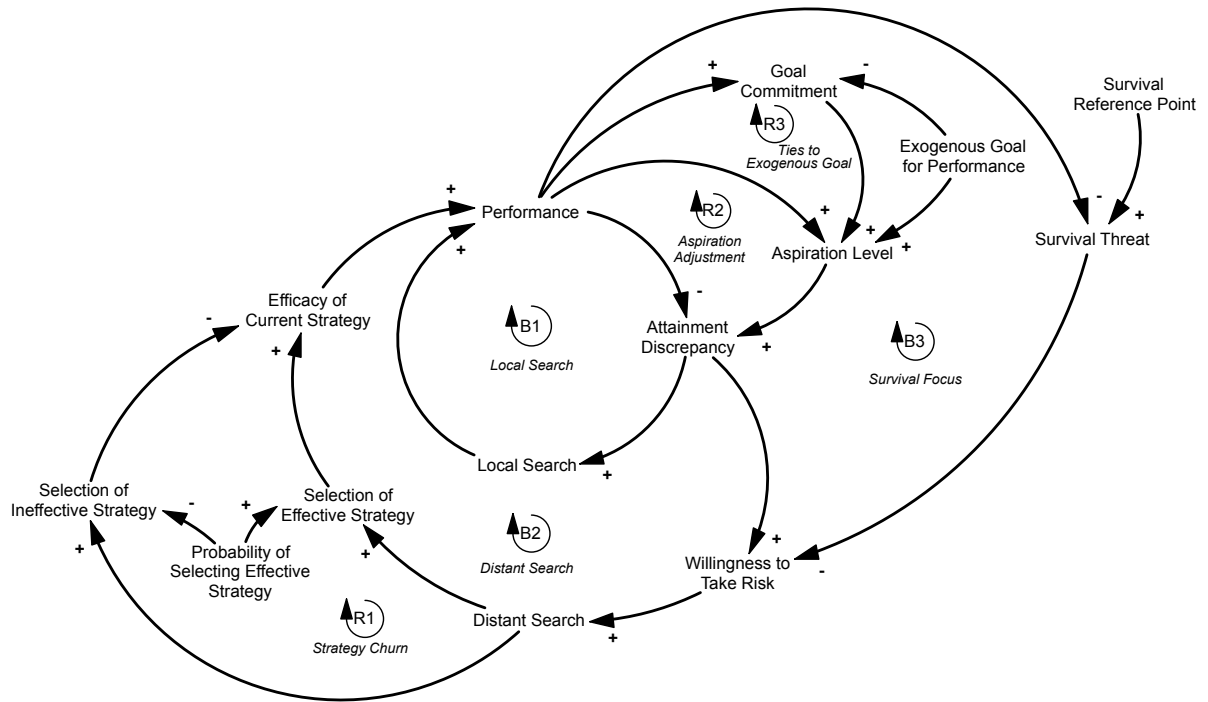


Figure 2 Illustration of hypothesized effects of stretch goals on performance level, variance, and skewness.

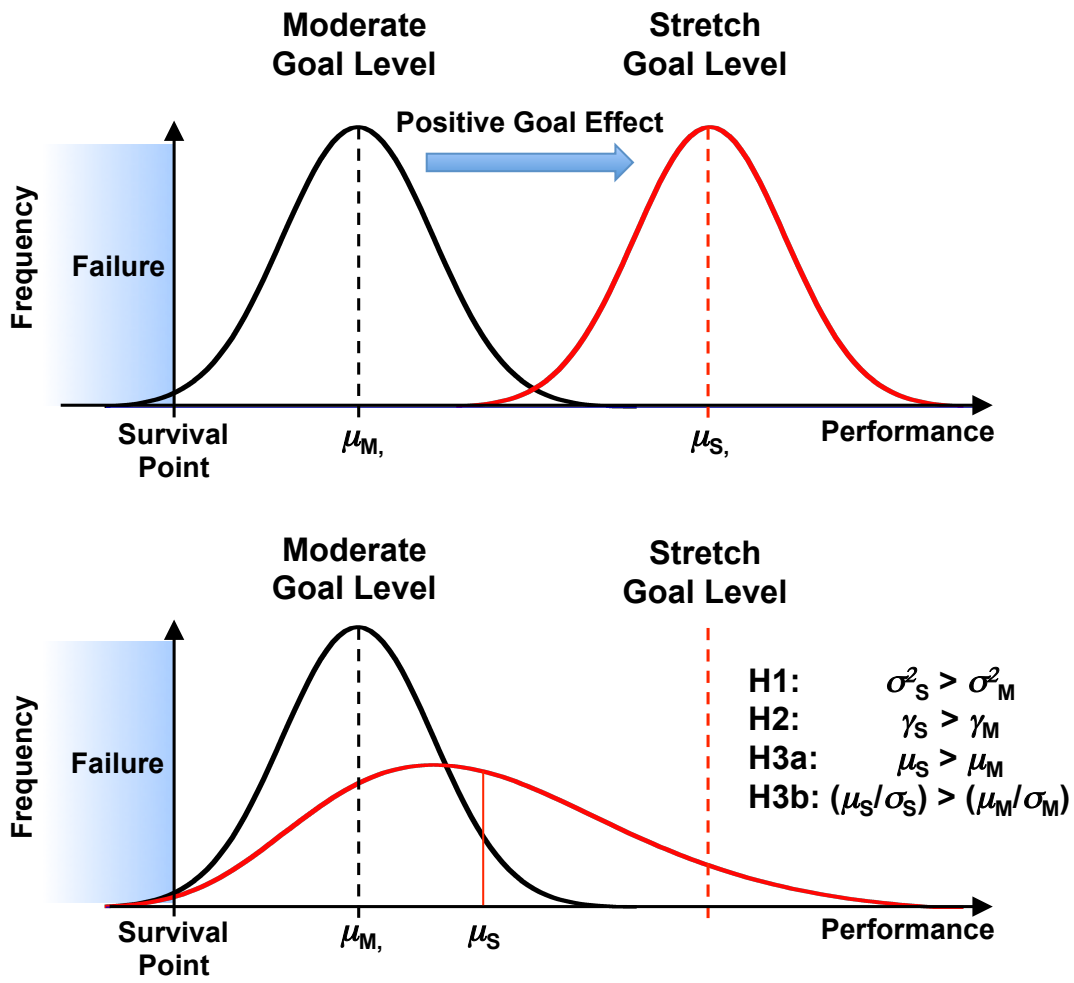


Figure 3 Panel A: Study 1 performance distribution at the end of Year 10 for Stretch and Moderate Goal Conditions

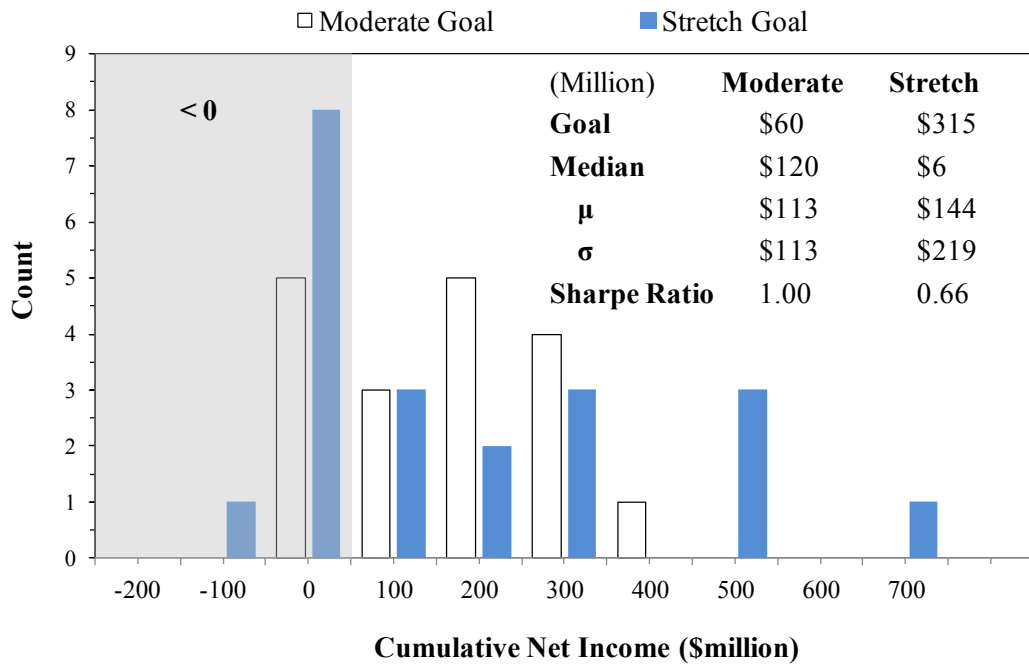


Figure 3 Panel B: Study 2 performance distribution at the end of Year 10 for Stretch and Moderate Goal Conditions

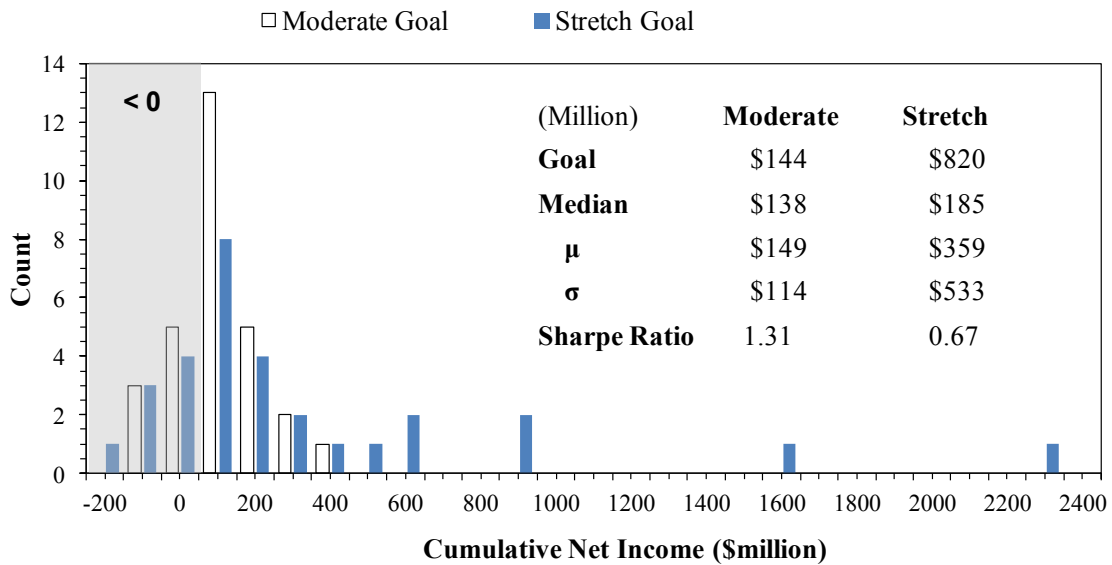


Figure 4 Mean Willingness to Take Risk (with error bars of +/- 2 standard errors) across the three simulation rounds

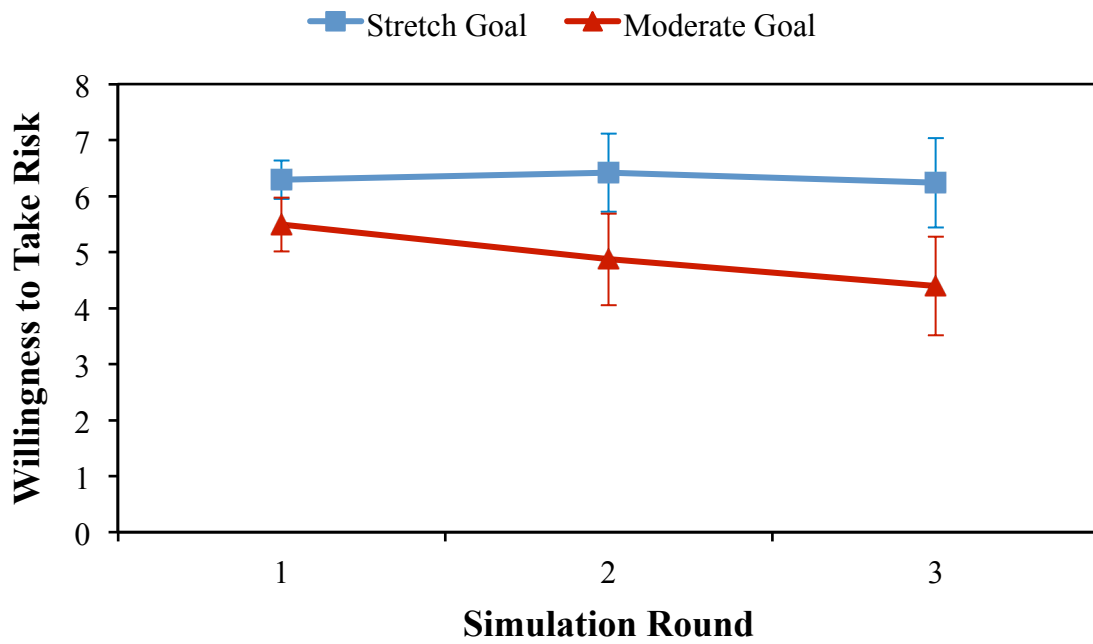


Figure 5 Mean Goal Commitment (with error bars of +/- 2 standard errors) across the three simulation rounds

