1.   Introduction

Goals have long played an important role in organization theory (e.g., Cyert and March 1963, Simon 1964) and are central to modern theories of organizational choice (Boyle and Shapira 2012, March and Shapira 1992, Sitkin et al. 2011). Goals influence how managers interpret organizational performance and frame their strategic responses (Lant and Shapira 2008, Lant 1992, Mezias et al. 2002, Short and Palmer 2003).

Over the last 40 years, many consultants, academics and authors have urged managers to adopt stretch goals in the belief that stretch goals motivate employees and generate high organizational performance. Collins and Porras (2002), for example, advocate the use of Big Hairy Audacious Goals (BHAG) to motivate, unify, and focus employee effort. Similarly, Thompson, Hochwarter, and Mathys (1997) recommend stretch targets to enhance motivation, performance, and creative decision making. Grant (2010) suggests top management should establish stretch targets to stimulate performance if the search for new strategies and new opportunities is limited by satisficing behavior. Many others have advocated the use of stretch goals (See, for example, Andrews 1971, Hofer and Schendel 1978, Peters and Waterman 1982, Slater 1999).

Following this advice, many organizations set stretch profit and growth goals (Fuller and Jensen 2010). For example, in 1990, Wal-Mart set stretch growth goals to double the number of stores and increase sales volume per square foot by 60 percent by the year 2000 (Collins and Porras 2002). General Electric set stretch goals in 1991 to improve operating margins by 36 percent and inventory turns by 100 percent within five years (Bartlett 2000, Slater 1999). Toyota Motors set stretch goals to reduce supply chain, distribution, and inventory costs by 50 percent from 1997-2001 (Oxnard 2004).

The accepted wisdom is that stretch goals improve organizational performance by disrupting complacency, promoting new ways of thinking, stimulating search and innovation, energizing employees, and guiding effort and persistence (for a recent review see Shinkle 2011). The implicit assumption is that stretch goals have a positive main effect on performance and that all organizations benefit from adopting stretch goals. However, Mosakowski (1998) raised questions about whether setting stretch organization goals is a simple 'rule for riches' for all organizations. Instead, she questions whether there is a more complex relationship between goal difficulty and organizational performance. In particular, Mosakowski theorized that more difficult, challenging goals increase performance variance across organizations, because high risk strategies lead to increasing pay-offs for positive outcomes at the expense of a higher probability of realizing lower performance outcomes.

Similarly, Sitkin et al. (2011) suggest that stretch goals do not always improve organizational performance. They discuss how stretch goals can generate positive performance effects for managers of high performing organizations with slack resources, but can generate negative effects for low performing organizations without slack resources. While Sitkin et al. focus on the effects of stretch

goals on the level of organizational performance, the patterns of behavior they describe would also increase the variance in performance across organizations.

These arguments imply that stretch goals benefit some organizations but not others. In addition, the organizational goal setting literature is silent on the effect on the shape of the organizational performance distribution when organizations adopt risky strategies to achieve stretch goals. Mosakowski (1998) emphasizes that understanding the impact of setting stretch goals requires investigating the shapes of distributions associated with alternative levels of goal difficulty.

Despite the widespread conventional wisdom encouraging managers to adopt stretch goals, there is limited empirical evidence about the effects of stretch goals on organizational performance. Therefore, the novel question guiding this research is:

> *What are the effects of stretch goals on the level of organizational performance, the variance in performance across organizations and the shape of the performance distribution?*

To begin to answer this question, we investigate the effects of stretch compared with moderate goals on performance in a laboratory experiment employing a widely used, realistic business simulation. Participants take the role of the CEO leading a start-up in a mature industry, specifically the airline industry, making strategic decisions to guide the organization in a competitive market environment over 10 years. We vary the difficulty of financial performance goals to examine how goal difficulty affects organizational performance.

The simulation presents decision makers with a complex, dynamic environment. Participants make five strategic decisions each quarter – aircraft acquisition, hiring, pricing, the range of services offered to customers, and marketing expenditure. They receive feedback on a range of operational, human resource, competitive, and financial outcomes. As in real businesses, the simulation includes multiple feedback processes linking employees, customers, competitors, the financial markets and other actors, with realistic time delays and nonlinearities in the impact of decisions on organizational performance. For example, in the long run hiring expands the firm's capacity to deliver high quality service, but in the short run recruiting and training new employees increases the workload of existing employees and dilutes the average experience of the workforce, potentially harming customer satisfaction.

The experiment enables us to examine the effects of stretch goals compared with moderate goals on performance. Of course, while experiments offer the ability to construct controlled settings, results from the laboratory are subject to an external validity threat. However, field studies are confounded by differences across organizations and circumstances, including differences in the nature and timing of performance objectives, in organizational resources and capabilities, and in competitive, end user, supplier, labor and capital market conditions, etc. The experimental design controls for these differences to isolate the effects of goal difficulty on organizational performance. The experimental results suggest where subsequent fieldwork should be focused to test the robustness of the results in naturalistic settings.

The paper is organized as follows. First, we review the relevant theory on organizational goals and develop three hypotheses for the effects on performance of stretch goals compared with moderate goals. We then describe the experimental procedure, the business simulator, and the results. To close, we review the validity threats to the findings and discuss implications for theory and practice.

2.   Theory and Hypotheses

Prior research shows that managers select organizational goals or aspiration levels using routines based on the organization's past goals and past performance, and the past performance of other comparable organizations (Cyert and March 1963: p. 115). Aspiration levels are anchored on past performance, and adjusted for what managers believe to be attainable improvements. For example, aspiration levels for organizational outcomes, including, for example, revenue, profit, or earnings per share, are often expressed as a target improvement over past performance (Fuller and Jensen 2010).

The literature provides strong support for the aspiration level adaptation model of managerial goal setting (e.g., Lant 1992, Mezias et al. 2002). That research explains how managers respond to attainment discrepancy gaps – the difference between goals and actual performance – by engaging in varying levels of search, learning, risk taking, and aspiration adjustment, including replacing aspirational with survival goals when they perceive risks of failure (Boyle and Shapira 2012, Bromiley 1991, Cyert and March 1963, Greve 1998, Lant and Shapira 2008, Lant 1992, March and Shapira 1987, Mezias et al. 2002).

Increasingly, as illustrated by the examples in the introduction, organizations have adopted stretch goals to replace incremental aspiration adjustments based on past performance. Changes in the external environment, the appointment of new senior leadership, pressure from investment analysts, or acquisition can trigger departures from traditional aspiration level adjustment routines and the adoption of stretch goals (Fuller and Jensen 2010). Similarly, in start-up firms, entrepreneurs frequently adopt stretch goals for growth and profit in the absence of established routines (Quindlen and Greenberg 2000).

We draw from the existing literature to motivate three hypotheses for the effects of stretch goals compared with moderate goals on organizational performance. We begin by examining the effect of stretch goals on the variance in, and skewness of, organizational performance. We then review the effect of stretch goals on the level of organizational performance, specifically, the goal main effect. The effects of stretch goals on the variance in, and skewness of, performance strongly condition the reasons for the presumed positive goal main effect on organizational performance.

2.1   Stretch Goals and Performance Variance

Recent research proposes that some organizations benefit from setting stretch goals, while others do not (Sitkin et al., 2011), suggesting that stretch goals compared with moderate goals lead to higher variance in performance among organizations. Previous research raised similar questions about

whether setting stretch goals is a simple 'rule for riches' for organizations, or if stretch goals result in high variance in performance among organizations (Mosakowski 1998). Here, we discuss the mechanisms connecting stretch goals and high variance in performance across organizations.

Adopting stretch goals compared with moderate goals creates larger attainment discrepancy gaps. Four managerial responses to discrepancy gaps discussed in the literature bear on the question of how goal difficulty affects the variance in organizational performance. The four responses involve (1) the impact of environmental complexity on search, (2) differences in managers' mental models, (3) goal commitment, and (4) risk taking.

First, larger attainment discrepancy gaps stimulate managers to engage in more intense search for and trials of new strategic options in their attempts to close the gaps between actual performance and their goals (Denrell 2008, Greve 2003). The search process generates a wide range of potential strategies with a variety of performance payoffs (Cohen 1984, Siggelkow and Rivkin 2006, Winter et al. 2007). The efficacy of search depends on the complexity of the environment, specifically, the costs of search relative to the expected yield, the time required to carry out search and to evaluate candidates for new strategies, and the reversibility of choices. Also, since the actual performance of a new strategic option cannot usually be evaluated until it is implemented, search efficacy improves with the quality of the mental models and data that managers use to estimate the expected performance of a strategy (Fang 2012).

As an example, consider a consumer goods company seeking to increase sales and profits for toothpaste. The costs and time delays in searching for superior product variants, packaging, pricing, and advertising are all relatively short. Market research can be carried out quickly and at moderate costs. Trials can be done in small regions and test markets, and are easily abandoned if ineffective. The environment favors effective search, rapid learning, and accurate mental models. In contrast, designing a new commercial aircraft such as the Boeing 787 or Airbus 380 involves enormous expense and long delays: the time and costs required for design, procurement and manufacturing must be borne before the first aircraft is available for testing. Unobtrusive trials in small test markets are impossible. The decision to go forward is not easily abandoned. Such complex environments reduce the efficacy of search routines. In those contexts, search consumes substantial financial, motivational and cognitive resources, frequently, reducing current performance. There are substantial delays in identifying new strategic options and implementing pilot projects to investigate the impacts. In such complex environments the additional search induced by stretch goals may enable some managers to find strategies that yield performance meeting or exceeding stretch goal levels. However, given the costs, delays and difficulty of evaluating the impact of strategies in advance, others will fail. In such complex settings search is likely to generate higher variation in strategies and performance for managers assigned stretch compared with moderate goals.

Second, intense search can lead to different rates of learning among managers with different levels of training, experience and ability to draw inferences from experience. As managers explore

the environment in response to attainment discrepancy gaps, they learn about the complex, dynamic decision environment in which they are operating by encoding inferences from experiences and developing mental models for interpreting that experience (Gary and Wood 2011, Levitt and March 1988). Learning in organizational environments is imperfect (Levinthal and March 1993, Paich and Sterman 1993). For example, Rahmandad, Repenning, and Sterman (2009) show that misperceptions of the delays between decisions and their impacts cause learning failure, even when all other conditions favor high performance and rapid learning (e.g., a simple task with a smooth, single-peak performance space with no measurement error or reporting delays).

For some managers, extensive search motivated by stretch goals facilitates the development of superior mental models as they continue to learn and improve. For those with poor mental models, learning is slow and search ineffective. For them stretch goals are a bridge too far (Janis and Mann 1977). Information processing becomes disorganized, inhibiting learning. As differences in managers' mental models, search, decision rules, policies and routines accumulate, differences in organizational performance grow, generating high variance in organizational performance.

Third, managers who face sustained and significant shortfalls in performance relative to a stretch goal vary in their commitment to those goals (Klein and Kim 1998). Some managers maintain high goal commitment and continue to pursue the stretch goals. Others become discouraged by failure decreasing their commitment to the stretch goals. Stretch goals are motivating for the former group of managers. They sustain positive affect, high self-efficacy and high commitment to the stretch goals, resulting in high performance (Sitkin et al. 2011). For the latter group, repeated failure to achieve stretch performance goals erodes self-efficacy and motivation, increases anxiety and stress, and reduces learning, goal commitment and performance. Stretch goals compared with moderate goals are likely to increase these differences in goal commitment among managers, increasing the variance in organizational performance.

Fourth, the higher attainment gaps caused by stretch goals compared with moderate goals result in higher risk taking. Models of managerial risk taking show that performance below a goal increases risk taking as long as performance is safely above the survival point (March and Shapira 1987). Similarly, prospect theory suggests that individuals performing below their reference point such as the goal are more willing to take risks than those performing above the reference point (Kahneman and Tversky 1979, Larrick et al. 2009), and risk taking has been found to be higher for organizations with performance below aspirations (Bromiley et al. 2001). High risk taking generates higher performance for some managers but lower performance for others. Therefore, stretch goals compared with moderate goals increase risk taking among managers, generating higher performance variance.

Summarizing the above arguments, stretch goals compared with moderate goals lead to higher attainment discrepancies, stimulating four managerial responses. Larger gaps drive greater search and variation in strategies, greater variation in learning and mental models, greater variation in goal commitment, and higher risk taking among managers. Taken together, we hypothesize that these

processes lead to higher performance variance for managers with stretch goals compared with moderate goals:

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

2.2   Survival Goals, Aspiration Adjustment, and the Performance Distribution

Advocates of stretch goals typically assume that formally assigned organizational goals become the operational goals guiding managerial behavior. Instead, managers frequently choose and construct the goals they adopt. These respond to a variety of factors, only one of which is the assigned goal. In searching for strategies to improve performance and achieve the stretch goals, some managers succeed and others fail. When organizational performance consistently falls short of the assigned goal, managers respond by shifting their attention to self-set, lower goals that depend on actual organizational performance (Lant 1992, March and Shapira 1987, Mezias et al. 2002). The result is that managers may focus on the assigned goal, a lower self-set performance goal, or a survival goal.

Managers in organizations achieving performance near or above the assigned goal are likely to stay focused on the assigned goal. They avoid taking risks that could reduce organizational performance below the goal. In the absence of disruptive changes in the environment, performance for this segment remains at or above the assigned goal. Typically, only a small proportion of organizations achieve their assigned stretch goals, which, by definition, are difficult to achieve.

However, managers in organizations that persistently perform substantially below the assigned goal often adjust their aspirations downward (Cyert and March 1963, Lant 1992, Mezias et al. 2002). The aspiration adjustment process reduces the attainment discrepancy gap. Instead of continuing to search and learn from performance feedback, these managers seek to enhance their self-image by assessing performance as satisfactory (Jordan and Audia 2012). Over time, current performance levels become the 'goal adopted in practice', replacing the assigned goals.

Other managers unintentionally adopt poor strategies or experience bad luck and face a threat to the survival of the organization (e.g., bankruptcy). Under such conditions, managers tend to replace the assigned goal with the goal of survival (Boyle and Shapira 2012, March and Shapira 1987, March and Shapira 1992). Shifting to a survival goal converts a large positive attainment discrepancy gap ('I'm falling far short of the assigned goal,') into a small negative gap, slightly above the survival goal ('I'm not dead yet,'). Attempting to survive, managers minimize risk taking. Local search dominates strategy development, learning is limited, and mental models become rigid (Bromiley 1991, Bromiley et al. 2001, March and Shapira 1992). Although some managers adopting a survival goal will go bankrupt, by avoiding risks and limiting the use of resources new strategies would require, others will succeed in avoiding failure. As a result, shifting to a survival goal constrains bankruptcies and truncates the left side of the performance distribution.

In summary, the internalized goals managers adopt frequently differ from their assigned goals. Some managers, faced with persistent attainment gaps, reduce the dissonance through goal erosion, while others, facing the risk of failure, adopt survival goals. The processes described above imply that stretch goals generate a right skewed performance distribution. Some organizations achieve the stretch goals, a few go bankrupt or hover near bankruptcy, and most perform below the stretch goal but remain viable.

Hypothesis 2: Compared with moderate goals, stretch goals generate right-skewed organizational performance distributions.

2.3   The Goal Main Effect on Organizational Performance

Advocates of stretch goals suggest that they improve performance by disrupting complacency, promoting new ways of thinking, and increasing energy levels throughout the organization (Hamel and Prahalad 1993, Rousseau 1997, Slater 1999, Thompson et al. 1997). A large body of work on goal setting in organizational psychology shows that specific, challenging goals increase average performance on simple, well-structured tasks (for reviews of this extensive research area see Locke and Latham 1990, Locke and Latham 2002, Locke et al. 1981). These findings support the common wisdom to adopt stretch organizational goals.

However, research also shows that more difficult goals can decrease performance on complex tasks with multiple alternative strategies (Earley et al. 1989, Wood et al. 1987). Research also shows that high goals are associated with dysfunctional strategic persistence following radical environmental change (Audia et al. 2000). A recent review article also argues that the beneficial effects of setting challenging goals have been overstated and that such goals can cause systematic harm (Ordóñez et al. 2009). In addition, some recent theoretical research argues that stretch goals reduce the performance of organizations that are already low performers with limited slack resources (Sitkin et al. 2011). Paradoxically, Sitkin et al. propose that these organizations are the most likely to adopt stretch goals in efforts to increase performance.

Overall, prior research findings are inconclusive about the effect of stretch goals compared with moderate goals on organizational performance. As a default option, we adopt the generally accepted hypothesis of a positive performance effect of stretch goals compared with moderate goals.

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

Figure 1 illustrates our hypothesized effects of stretch compared with moderate goals on organizational performance, the variance in performance across organizations, and the shape of the performance distribution. The top part of the figure shows a hypothetical distribution of performance under moderate goals as a normal distribution and a potential distribution under stretch goals with the same variance and skewness, and performance shifted to the right. This is the implicit assumption of how stretch goals improve performance. The bottom part of Figure 1 shows the same hypothetical

distribution for moderate goals and an alternative distribution under stretch goals with higher variance, increased right-skewness, and higher median performance that illustrates our three hypotheses.

Insert Figure 1
----

3.   Methodology

We use a laboratory experiment based on a simulated organization to isolate the effects of goal difficulty on performance. Specifically, we compare the effects on performance of stretch and moderate goals achieved by participants who manage the simulated organizations. Participants were rewarded for achieving the assigned goals through pay-for-performance incentives.

3.1   The Simulated Organization

The People Express simulation is an interactive, computer-based management simulation of an airline operating in a competitive market. The simulation captures many well-established features of competition between new entrants and incumbents in the airline industry. Participants take on the role of the Chief Executive Officer (CEO) of a start-up airline, making quarterly decisions for aircraft orders, employee hiring, average fare, marketing expenditure and service scope. Their goals are to deliver on stretch or moderate profit goals over a forty-quarter simulation.

The management 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, nonlinear relationships, and accumulations (Graham et al. 1992). These features are common in the complex environments that managers operate in when making strategic decisions. The simulation has also been utilized in previous research (Bakken et al. 1992, Graham et al. 1992), and is used frequently in MBA and executive teaching.

Growing the simulated airline from startup to major carrier involves a number of trade-offs. For example, setting low fares increases customer demand. Rising demand can be met by expanding the fleet. However, meeting rising demand also requires employees who can provide high quality service, including, for example, reservations and bookings, check-in, boarding and de-boarding of aircraft, error-free baggage handling, catering and a range of in-flight services.

It takes time to expand both the number and skills of employees. Hiring and training delays can result in staffing shortages and skill dilution as decision makers seek to grow their fleet and customer base. Staff and skill shortfalls increase the workload, leading to long hours, fatigue and burnout. Fatigue leads to higher employee turnover, low morale, and lower service quality. Rising employee turnover further increases workloads and working hours for the remaining employees, and increases the time experienced employees must spend recruiting and training new employees. New recruits are

less skilled than experienced employees and, as the proportion of new recruits increases, average service quality decreases.

Declining service quality undermines the reputation of the business in the marketplace, eroding demand growth and harming capacity utilization and financial performance, with potentially irreversible outcomes. If decision makers successfully navigate the above trade-offs and grow the airline, competitors respond with aggressive pricing in an attempt to limit their market share loss. Despite these difficulties, it is possible to grow the airline profitably, if decision makers adopt an appropriate strategy. Consistent with the simulation, the majority of low-cost, no-frills airlines have failed, both in the US and Europe (Button 2012). A few have survived to date (e.g., RyanAir, Spirit) and still fewer have profitably achieved scale (e.g. Southwest Airlines).

3.2   Laboratory Study

Participants: Fifty-nine students from a large university participated in the study. Participants averaged 22 years of age, with 53% female. Eighteen percent of participants were majoring in economics, 24 percent were majoring in management – including marketing, accounting, finance, strategy, IT management, and organizational behavior – with the others majoring in a broad range of other fields. Individual participants were randomly assigned to achieve stretch (n = 30) or moderate (n = 29) profit goals.

Procedure: Each participant followed instructions on an individual computer and managed their simulated firm independently. At the start of a session, the experimenter read the instructions aloud, and participants then spent 25 minutes working through a set of introductory exercises to become familiar with the simulation[1].

After completing the introductory exercises, each participant received a brief outlining the performance goals for their simulated firm. Consistent with research that shows organizations primarily use financial accounting measures to assess performance (Short and Palmer 2003), we use cumulative profit as the organizational performance measure and set goals for the level of cumulative profit to attain in each year. Participants received a table showing their target for cumulative profit (See Tables 1 and 2 for the moderate goal or stretch goal, respectively). Targets for cumulative profit are assigned for years 3 through 10 of the simulation. No goal is assigned through year 2 because startups such as those represented by the simulation often initially lose money before becoming profitable. Subjects were not informed whether they received the moderate or stretch goals.

After reading the brief, participants answered a five-item questionnaire to assess whether they were committed to achieving the assigned goals. Research shows that commitment to goals plays a critical role in determining goal effects on performance (Hollenbeck et al. 1989, Klein et al. 2001, Locke et al. 1988). The five items were adapted from items that have been used extensively in prior

---

[1] The set of exercises has been submitted as a supplementary document and could be made available online.

research (Klein et al. 2001). An example item was: "Quite frankly, I don't care if I achieve the annual goals or not." For each item, participants recorded their level of agreement on a ten-point scale where "strongly disagree" is coded 0 and "strongly agree" is coded 10. Goal commitment was assessed by the average across the five items (after reverse coding some items to align the scale). A high score on the scale indicates a high level of goal commitment. The measure was collected before each of the three simulation rounds.

After answering the goal commitment questions, participants completed the first round of managing their simulated firm (Simulation Round 1). The round lasts 40 quarters, ending earlier if the participant's decisions lead to bankruptcy.

Insert Table 1 and Table 2

After all participants completed the first simulation round, the same procedure was repeated for the second and third rounds. The first two rounds were treated as a learning phase. Therefore, prior to the third round testing phase, decision makers potentially had 80 decision trials, two complete simulation rounds of 40 quarters each, or twenty years of simulated experience, to learn about the management simulation.

Goal Manipulation: Goal levels were selected to represent moderate and stretch goals for cumulative profit for years 3 through 10 of the simulation. The average compound growth rates of cumulative profit in the moderate and stretch goal conditions are 29%/year and 47% year respectively. Although difficult, the stretch goal levels are achievable.

We examined the performance of numerous benchmark strategies for managing the simulated firm. The benchmark strategies consist of simple decision rules for price setting, fleet acquisition, hiring, etc. The benchmark rules are not optimal, but constitute behaviorally realistic, boundedly rational heuristics such as 'hire enough people to replace employee attrition plus a certain number for each new aircraft acquired.' Despite being less than optimal, the performance of the benchmark decision rules exceed the stretch goal. For example, a plausible set of benchmark decision rules generates cumulative profit in year ten 5% higher than the stretch goal for that year. More sophisticated (but still behaviorally plausible) strategies yield substantially higher performance, for example 265% higher than the stretch goal for year 10.

Participants were paid $5 for participating in the experiment plus an additional cash payment of $2 for each intermediate year (3-9) in which their cumulative profit met or exceeded the target for that year, and a payment of $6 if their cumulative profit in year 10 met or exceeded the final goal (Tables 1 and 2). Participants attaining their goals in all years of the three trials would earn up to $65 (a maximum of $20 in each of the three rounds of the simulation, plus the $5 participation payment).

4. Results

Hypotheses 1-3 were tested using the performance data for the third simulation round, after the two learning rounds. For those participants whose firms go bankrupt, we use cumulative profit from the quarter in which the simulated organization goes bankrupt as the measure of performance for that year. Participants were assigned specific goals for years 3 through 10 of the simulation, and Hypotheses 1-3 are tested using these repeated measures, with 59 observations in each of the eight years for a total of 472 observations. Each individual decision maker's performance is subject to auto-correlation across repeated measures, so all tests are either adjusted for multiple dependent hypotheses testing or the correlation structure is specified in the estimation procedure.

Hypothesis 1 is examined with the Fligner-Killeen test for homogeneity of variances, using the Holm-Bonferroni adjustment for multiple dependent hypotheses testing. The Fligner-Killeen test is robust against departures from normality (Conover et al. 1981). In addition, the more familiar Levene test for equality of variances is reported, using the Holm-Bonferroni adjustment for multiple dependent hypotheses testing. Both the Fligner-Killeen and Levene tests assess whether there are differences in the variance in performance between the stretch and moderate goal groups each year for Years 3 through 10 of the simulation.

To test Hypothesis 2, we calculate the sample skewness parameter for both the moderate and stretch goal conditions. For a sample of 30 observations the critical value for $p < 0.01$ is 1.17. Skewness parameters are calculated each year for Years 3 through 10 of the simulation, correcting the test p-values for multiple dependent hypotheses testing using the Holm-Bonferroni adjustment.

Hypothesis 3 is subject to three tests. First, we apply the Mann-Whitney nonparametric test for Years 3 through 10, applying the Holm-Bonferroni adjustment for multiple dependent hypotheses testing. Under the null hypothesis, the distributions of both groups are equal. For continuous response variables, as in this case, a significant Mann-Whitney test can be interpreted as showing a significant difference in medians.

Second, as a robustness test, a median regression model was also estimated for Year 10, with Goal Condition as an independent variable and standard errors computed over 2,000 bootstrapping samples. Third, Hypothesis 3 is tested using a χ^2 independence test of the frequency distribution for organizations classified as either below or above the final (year 10) moderate goal performance level.

Figure 2 shows the performance distributions at the end of Year 10 for moderate and stretch goals. Performance in the moderate goal condition does not significantly depart from a normal distribution (Kolmogorov–Smirnov D[29] = .16, $p = .06$). However, performance in the stretch goal condition is not normally distributed (Kolmogorov–Smirnov D[30] = .25, $p < .001$). Stretch compared with moderate goals result in a greater range of performance outcomes, with both higher and lower performance.

In the moderate goal condition, 17% of participants go bankrupt; 41% achieve or exceed the long run, Year 10, profit goal; and the remaining 42% avoid bankruptcy but do not successfully achieve the long run profit goal. For decision makers assigned stretch goals, 23% go bankrupt; 13% achieve or exceed the long run profit goal; and the remaining 64% avoid bankruptcy but do not successfully achieve the long run goal.

Insert Figure 2 here

As a manipulation check of whether the stretch and moderate goals were of substantively different levels of difficulty, we test whether the attainment discrepancy gap and the number of goals achieved over the ten years are significantly different between the two conditions. Decision makers assigned stretch goals experience significantly higher discrepancy gaps (Mann-Whitney's $U = 44.00$, $z = -5.93$, $p < 0.01$) between the assigned goal and their actual performance (median discrepancy gap = $694.0 million) compared with subjects in the moderate goal group (median discrepancy gap = $9.4 million).

In addition, decision makers assigned moderate goals achieve a higher number of goal levels than decision makers who are assigned stretch goals. The former met an average of 5.2 out of their eight goals. The latter met an average of 1.1 out of a total of eight profit goals. The difference is significant (Mann-Whitney $U = 139.5$, $z = -4.66$, $p < 0.000$).

Decision makers assigned stretch goals rarely achieve the assigned goals and usually generate large attainment gaps between actual performance and the stretch goal. The stretch goals compared with moderate goals are significantly more difficult for participants to achieve.

4.1   Tests of Hypotheses

Hypothesis 1, *Stretch compared with moderate organizational goals generate higher variance in organizational performance,* is supported. Table 3 reports the goal levels, median cumulative profits, and standard deviations for both goal conditions over the ten year simulation. Table 3 also reports the Fligner-Killeen and Levene test statistics for evaluating differences in performance variance between the stretch and moderate goal groups. All decision makers are given the same initial conditions in Year 0 and the assigned goals begin at the end of Year 3. It takes time for the effects of different goal difficulty levels and the associated managerial responses to affect performance. The effects of stretch compared with moderate goals result in significantly higher variance in organizational performance from Year 6 onwards (all Holm-Bonferroni adjusted p-values for Fligner-Killeen and Levene tests < 0.05 for Years 6 through 10)[2]. For example, at the end of Year 10, decision makers assigned stretch

---

[2] The Levene test results are robust when computed using alternative location estimators, including replacing the mean in Levene's formula with the median or with the 10% trimmed mean.

goals compared with moderate goals exhibit significantly higher performance variance (SD = $526.6 million versus $114.5 million, respectively; L[1, 57] = 13.15, Holm-Bonferroni adjusted p = .005).

Insert Table 3 here

Hypothesis 2, *Stretch organizational goals generate right-skewed organizational performance distributions,* is also supported. From Year 7 onwards, the distribution of performance outcomes under stretch goals is significantly right-skewed (all Holm-Bonferroni adjusted p-values for skewness parameter tests < 0.01 for Years 7 through 10). For example, at the end of Year 10, the skewness parameter for the stretch goal group is 2.38, compared with the critical value of 1.17 (Holm-Bonferroni adjusted p < .001). Figure 2 illustrates clearly the right-skewed distribution of performance for the stretch goal condition. In contrast, the distribution of performance under moderate goals is not skewed (all Holm-Bonferroni adjusted p-values are not significant).

Hypotheses 3, *Stretch compared with moderate organizational goals increase organizational performance*, receives mixed support. First, applying Mann-Whitney tests, there is no significant difference in cumulative profits between the stretch and moderate goal conditions (all Holm-Bonferroni adjusted p-values are not significant for Years 3 through 10). For example, at the end of Year 10, the difference in cumulative profits of decision makers assigned stretch goals (Median = $185 million) compared to moderate goals (Median = $138 million) is not significant (Mann-Whitney's U = 348.0, z = -1.32, Holm-Bonferroni adjusted p = 0.68, r = 0.17). Nevertheless, the effect size (r = 0.17) indicates that there is a small difference in cumulative profits between stretch and moderate goals.

Second, median (quantile) regression estimates for the effect of Goal Condition on median performance in Year 10 show the above Mann-Whitney tests results are robust ($\beta_{Goal\_Condition}$ = 55.91, bootstrapped S.E. = 47.96, p = 0.25, 95% Conf. Int. [-40.14, 151.97]). The pseudo R^2 for the overall model = 0.02, indicating Goal Condition is not a good predictor of median performance[3].

Third, we tested whether stretch compared with moderate goals resulted in a higher frequency of organizations achieving or exceeding the long-run (year 10) moderate goal level (see Table 4). The frequency of performance equal to or above the moderate goal is significantly higher for decision makers assigned stretch goals compared with those assigned moderate goals (χ^2 = 4.06, p = .04). More decision makers in the stretch goal condition meet or exceed the long-run moderate goal performance level than those assigned moderate goals.

Insert Table 4 here

_____

[3] The model estimated was: **Performance = Constant + β(Goal Condition) + ε.**

Overall, these results for the effect of stretch goals compared with moderate goals on the level of performance are mixed. Although stretch goals do not lead to significantly higher median performance, there is a small positive effect ($r = 0.17$) and decision makers assigned stretch goals more frequently meet or exceed the long-term moderate goal compared with decision makers assigned moderate goals. Inspecting Figure 1, the effect of stretch goals compared with moderate goals on the level of performance is small compared to the effects on performance variance and skewness. Stretch goals lead to high performance for a small proportion of organizations that achieve (or nearly achieve) the stretch goals. That improvement comes at a cost, however, as a higher proportion of participants in the stretch goal condition go bankrupt (23% vs. 17% in the moderate goal condition). Given these mixed results, we conclude that the positive goal main effect on organizational performance, if any, is small and dominated by the effects of stretch goals on performance variance and skewness.

4.2   Analysis of Subject Behavior

To further explore the effects of stretch goals, we report on commitment to the assigned goals for the stretch and moderate goal conditions. As shown in Figure 3, both groups start simulation round one with the same level of goal commitment, but goal commitment declines in simulation rounds two and three for those assigned stretch goals. Prior to the final simulation round, decision makers assigned stretch goals report a mean goal commitment level of 3.99 on a ten-point scale, compared with a mean of 6.12 for participants assigned moderate goals. Consistent with higher attainment discrepancy gaps, decision makers assigned stretch goals compared with moderate goals exhibit significantly lower goal commitment ($F[1, 57] = 17.49$, $p < 0.001$), increasing the probability of downward aspiration level adjustments.

Insert Figure 3

While goal commitment in the stretch goal condition declines, aircraft acquisition and employee hiring are higher than for the moderate goal condition. By the end of Year 10 in the final simulation round decision makers assigned stretch goals expand the fleet to a median level of nine aircraft with over 665 total employees. Decision makers in the moderate goal condition expand the fleet to a median level of four aircraft with 300 total employees. Participants assigned stretch goals compared with moderate goals acquire significantly more aircraft (Mann-Whitney's $U = 276.0$, $z = -2.42$, $p < 0.05$) and hire significantly more employees (Mann-Whitney's $U = 273.00$, $z = -2.46$, $p < 0.05$).

Decision makers assigned stretch goals adopt and execute strategies favoring faster growth than those assigned moderate goals. However, as the tests of Hypothesis 3 show, the median or typical organization assigned stretch goals did not grow the aircraft fleet rapidly nor profitably enough to achieve the stretch profit goals within the timeframe.

Expanding the fleet and hiring employees to maintain the service capacity of the airline are related. Without adequate staffing, as the fleet grows, service quality suffers, driving passengers away and leading to low load factors and large financial losses. Building a skilled workforce is further complicated by the delays in hiring and training, and the impact of inexperienced employees on service quality, workload, burnout and turnover. This coordination challenge often undermines decision makers' attempts to grow the aircraft fleet rapidly in the simulation (Graham et al. 1992). Frequently, after a burst of fleet growth, service quality declines, load factors drop, and growth in the fleet must be halted or the airline would go bankrupt (Graham et al. 1992, Sterman 1988).

Top Performers in Stretch Goal Condition. To explore the differences in aircraft purchasing and employee hiring decision making among participants assigned stretch goals, we compare the top quintile based on cumulative profit at the end of Year 10 with the remaining 80 percent of participants. The top 20 percent expanded to a median fleet of 40 aircraft and 2,709 employees compared with 6.5 planes and 464 employees for the others (Mann-Whitney's $U = 10.50$, $z = -3.20$, $p < 0.001$ for aircraft and $U = 10.00$, $z = -3.22$, $p < 0.001$ for employees). Decision makers in the top 20 percent of the stretch goal condition successfully identified strategies that allowed them to profitably coordinate the growth of the fleet and employees. The other 80 percent of decision makers failed to do so.

5.   Discussion

The findings above are discussed here under three headings: variance in performance, skewness and the goal main effect. Next we consider broader implications for theory and practice.

5.1   Variance in Organizational Performance

Some scholars have questioned the conventional wisdom that stretch goals are beneficial suggesting that stretch goals benefit some organizations but not others. To our knowledge, this is the first study to test these theories. Sitkin et al. (2011), for example, posit that stretch goals benefit only high performing organizations with slack resources, but undermine the performance of low performing organizations without slack resources. These effects would lead to higher variance in organizational performance for stretch goals than for moderate goals, as seen here. In addition, the findings support prior research asserting that difficult organizational goals increase performance variance compared with less difficult goals (Mosakowski 1998).

Mosakowski (1998) based her expectation about the relationship between difficult organizational goals and performance variance on findings from organizational psychology showing goal difficulty increases performance variance on well-structured tasks (Erez and Zidon 1984, Locke and Latham 1990, Locke 1982). Our findings extend that research by showing stretch compared with moderate organizational goals lead to higher performance variance in more realistic, less structured and more complex tasks such as managing a firm. Together, these findings establish variance in performance

among organizations as an important phenomenon and should not be treated as merely error around the presumed goal main effect. Stretch goals are not a simple "rule for riches" and not all organizations are likely to benefit from adopting stretch goals.

We find, as expected, that stretch organizational goals generate larger attainment discrepancy gaps than moderate goals. This is the mechanism by which stretch goals are intended to motivate decision makers to strive for higher performance (Shinkle 2011, Sitkin et al. 2011). Prior research shows that larger discrepancy gaps result in more extensive search (Denrell 2008, Greve 2003, Lant and Montgomery 1987). Previous research also shows that large attainment discrepancy gaps increase decision makers' willingness to take risks to achieve their goals (Bromiley et al. 2001, Greve 1998) as long as the organization is not near the survival point (March and Shapira 1987, March and Shapira 1992). For example, Greve (1998) shows organizations that perform below social (i.e., industry average performance) and historical (i.e., prior actual performance) aspiration levels engage in more risky organizational changes. In addition, recent judgment and decision making research shows that challenging goals compared with "do your best" goals increase risk taking in negotiations (Larrick et al. 2009). Our results are consistent with these prior studies examining the effects of attainment discrepancy gaps on organizational search and risk taking, and extend that research by integrating large attainment gaps with performance variance across organizations. Larger discrepancy gaps induce more extensive search and risk taking. These result in higher performance variance because some managers discover and implement effective strategies while others do not.

Decision makers assigned stretch goals compared with moderate goals adopt higher asset growth strategies. Specifically, in this study, participants assigned stretch goals compared with moderate goals, purchase more aircraft and hire more employees. However, over 80 percent of decision makers assigned stretch goals fail to achieve the long run (i.e., Year 10) profit goal because they do not succeed in coordinating increases in the aircraft fleet with hiring enough employees to achieve the stretch profit goals. While strategies to grow assets are relatively simple to identify and implement, profitable growth strategies to achieve stretch profit goals in dynamic competitive environments are complex and hard to discover. These results are consistent with prior research showing widespread misperceptions of feedback. Inaccurate and incomplete cognitive maps of causal relationships and limited ability to mentally simulate the dynamics of complex systems hinder attempts to improve performance (Gary and Wood 2011, Paich and Sterman 1993, Sterman 1989).

The supplementary analysis shows that the top performing 20 percent of decision makers in the stretch goal condition adopt and successfully execute profitable asset growth strategies. Stretch goals motivate these decision makers to identify profitable growth strategies to achieve (or nearly achieve) the profit goals. These findings are consistent with prior work examining variation in mental models and the link to performance (Gary and Wood 2011), and extend that research by highlighting how goals affect the heterogeneity of mental models and strategy choices.

5.2   Right-Skewed Performance Distribution

Stretch goals generate significantly right-skewed performance distributions. These findings identify the impact of stretch goals on the shape of the performance distribution and extend prior research asserting that setting stretch organizational goals represents a "rule for chances of riches" (Mosakowski 1998). Decision makers respond in diverse ways to stretch goals, partitioning organizations into three performance levels.

First, there are the managers for whom stretch goals motivate strategy search that leads to high firm performance. These top performers identify and successfully execute profitable asset growth strategies. The top 20 percent accounted for 63 percent of the total cumulative profits earned in the stretch goal condition. The finding that only a small proportion of decision makers achieve the stretch goals is consistent with prior research showing dynamic complexity – time delays, feedback effects and nonlinearities – undermines decision making and performance (Diehl and Sterman 1995, Sterman 1989, Sterman 1989).

Second, there are managers for whom both moderate and stretch goals motivate strategy search that leads to bankruptcy or poor firm performance. Presumably these managers, even those with low goal commitment, do not seek to fail. Since there are no random events in the simulation, poor performers and those who go bankrupt must either have a poor mental model that leads them to adopt a strategy that threatens survival or find a reasonable strategy but fail to implement it consistently or successfully. Prior research suggests these managers, facing the risk of failure, are likely to substitute a survival goal for the initial stretch goal (March and Shapira 1987, March and Shapira 1992). Shifting to a survival goal curtails search and risk taking, and increases the likelihood of remaining near the survival point.

Third, many managers assigned stretch goals fail to identify profitable growth strategies to achieve the stretch profit goals but do avoid bankruptcy. The results show 64 percent of decision makers assigned stretch goals fall into this category. These managers do not identify profitable growth strategies. Instead, decision makers in this performance region abandon growth strategies and, in so doing, abandon any chance of achieving the stretch profit goals. Given the difficulty of identifying profitable growth strategies in dynamically complex environments (Sterman 1989, Sterman 1994), it is not surprising that the majority of decision makers end up in this performance region. Prior research shows decision makers engage in aspiration level adaptation to reduce attainment discrepancy gaps (Lant 1992, Mezias et al. 2002). Our results are consistent with this goal erosion process, with these managers reporting low goal commitment.

Overall, the combined effects of stretch goals to partition organizations into these three categories explains the skewed distribution of organizational performance. The results build on prior research on aspiration level adjustments (Lant 1992, Mezias et al. 2002), survival reference points (Boyle and Shapira 2012, March and Shapira 1987, March and Shapira 1992), and search (Denrell 2008, Greve

1998, Greve 2003) to explain how stretch goals affect the overall shape of the performance distribution.

5.3   The Goal Main Effect

Stretch goals compared with moderate goals do not lead to higher performance for the median organization. However, stretch goals increase the proportion of organizations that equaled or exceeded the moderate goal. These findings are consistent with prior research showing mixed results regarding the performance effects of goal setting on complex tasks (Chesney and Locke 1991, Earley et al. 1989, Larrick et al. 2009, Smith et al. 1990, Wood et al. 1987). The results are also consistent with recent research proposing that some organizations benefit from setting stretch goals while others do not (Sitkin et al., 2011).

Here, stretch goals enabled a small proportion of organizations (13 percent) to achieve high performance. At the other end of the spectrum, 23 percent of the organizations go bankrupt in response to stretch goals. The majority of performance outcomes resulting from stretch goals are above bankruptcy but below the assigned goal. There is no significant difference in median performance between the stretch and moderate goal conditions. In contrast to the accepted assumption of a positive goal main effect, there is no evidence of a strong positive goal main effect of stretch compared with moderate goals on organizational performance. While challenging compared with do-your-best goals motivate problem solvers to achieve higher performance outcomes on simple, highly-structured tasks (Locke and Latham 1990), stretch organizational goals enable only a small proportion of decision makers to achieve high organizational performance in dynamically complex environments.

Partitioning the goal effect into a goal motivation and a strategy efficacy effect helps explain the different findings for the goal main effect in the psychology literature and this study. In well-structured problems, where all actors understand the strategy to follow, the primary driver of performance is motivation. So, the goal motivation effect increases performance for most actors. In contrast, in complex problems in dynamic environments, the primary driver of performance is the ability to identify a successful strategy (Wood and Locke 1990). The motivation to search for such a strategy is a necessary but not sufficient condition for success. Instead, the capabilities to identify a successful strategy are critical and not equally distributed across actors.

5.4   General Research Implications

Our findings highlight the importance of examining the effects of different goal levels on performance variance. Typically, strategy and organization theory research restricts its analysis to investigating

differences in levels of performance associated with different strategic and organizational choices[4]. As a result, there is limited knowledge about how different strategic choices – including setting different goal levels – affect the distribution of performance across actors (Mosakowski 1998).

The results also raise questions about explaining organizational performance as a function of decision makers' motivation and, more importantly, of their imputed motivation. Initially, at the start of the first simulation round, the two groups of participants were equally motivated to achieve their assigned goals. For those that failed to achieve the stretch goals, the explanation lies in their failure to identify successful strategies. Poor performance is a consequence of their inadequate strategic analysis and not of their initial motivation.

5.5   Limitations

Three choices were made that contribute to the external validity of our findings. First, dynamic decision making experiments using complex management simulations capture many aspects of the decision making environments of managers, including feedback, time delays, accumulations and nonlinearities. There is widespread evidence that characteristics of the decision environment – particularly complexity – affect psychological processes and empirical results (Gary and Wood 2011, Paich and Sterman 1993). The management simulation used in this study has been calibrated to represent realistic decision making challenges faced when launching a start-up airline to compete against entrenched incumbents (Graham et al. 1992).

Second, prior research on organizational goals and aspirations has typically inferred goal levels based on each firm's prior performance or industry average performance (Miller and Chen 2004), rather than measuring goals directly. The experimental design adopted here allowed for the direct manipulation of goal difficulty levels to directly test the hypothesized causal relationships. This would be difficult, if not impossible, in the field.

Third, examining the impact of organizational goal difficulty on performance through behavioral experiments has the advantage of controlling for between-industry differences, such as growth rates, cyclicality, capital intensity, regulation, and concentration or market structure. Future research can build on this work by testing the generalizability of the findings reported here both in the field and in laboratory experiments across a variety of market contexts.

However, there are also limitations and potential validity threats to the findings discussed here. First, this study examines the behavior of individual decision makers managing a simulated organization and does not explore the multi-faceted socio-political process of goal setting and decision making in organizations. Different organizational coalitions may have diverse views about goals, leading to conflict, adoption of different sub-goals, and coordination problems (Chattopadhyay

---

[4] An exception is the work on risk taking examining the relationship between longitudinal variance in firm returns and performance (Bromley et al. 2001.

et al. 1999, Cyert and March 1963). Research should address these issues in a range of field settings to enhance external validity. However, such studies would be challenging and strong results difficult to tease out from the many sources of uncontrolled variation, exogenous and endogenous, that arise in field settings. Experimental studies, though simplified, offer the ability to control conditions and help identify those factors and processes that appear to be most important, and can be seen as tools to help focus field work on those issues likely to be most critical.

Second, novices rather than experienced managers made strategic decisions for the simulated organizations in this experiment. Managers leading organizations often draw on a large base of experience and substantial domain-specific knowledge to make decisions. Nonetheless, the relationship between goals for overall organizational performance and actual performance is complex. Regardless of how much experience managers may have, they do not know with certainty what impact their strategic choices will have on subsequent firm performance (Mosakowski 1998). Also, research shows there are no strong effects for education, age, or prior work experience on decision makers' ability to understand dynamic complexity (Cronin et al. 2009, Sterman 2010).

Third, the performance goals participants faced were not adaptive. The specified stretch goals were ambitious, though achievable, and did not erode in the face of poor performance; a participant who fell behind early in the simulation faced a tougher task than one who did well early on, perhaps contributing to the decline in goal commitment and performance. In many organizations, however, goals are adaptive so that poor performance does not lead to attainment gaps so large that people become discouraged. For example, percentage increases over current performance at any point in time would maintain constant goal difficulty levels and may be more in line with how most organizations set goals. A large expected improvement over current performance would be more of a "stretch" than a small increment but would not punish poor performance. Future work should investigate the effects of adaptive goals, though it is unclear whether such flexible goals are consistent with the idea of the Big Hairy Audacious Goal.

Fourth, we assigned decision makers just one goal dimension, cumulative profit, in our experiment. However, managers adopt, are assigned or are rewarded based on a wide variety of different goals, including, for example, profit, growth, share price and market share (Short and Palmer 2003). In addition, in our experimental study, we tested the effects on performance of only two goal levels, stretch and moderate goals. Other levels of goals could be tested. The definitions of stretch goals and gradations of more or less difficult goal levels are not well-defined in the literature. Future research should examine the effects of pursuing multiple goal dimensions simultaneously and investigate whether the relationships between goals and performance change as managers move from easy-to-achieve goals to seemingly impossible ones.

Fifth, stretch goals may have unintended side effects including unethical behavior, corrosion of organizational culture, and illegal activity (Mishina et al. 2010, Ordóñez et al. 2009). Examples

include Enron, Satyam, WorldCom, and other corporate scandals. These issues are not investigated in the simulation, limiting the negative effects that stretch goals might activate.

5.6   Implications for Practice

Popular managerial writings overwhelmingly favor the use of stretch goals to improve organizational performance (Hamel and Prahalad 1993, Peters and Waterman 1982, Slater 1999). Managers in publicly listed companies have increasingly embraced the practice of announcing stretch financial performance goals (Fuller and Jensen 2010). These stretch goals often take the form of targets for annual or quarterly growth in earnings, earnings per share, stock price, revenue, market share, or similar performance metrics.

The results of this study present a complex picture of the effects of goal difficulty on the distribution of and variance in performance. Where the advice to adopt stretch organizational goals is based on the assumption of a strong positive main effect of goal difficulty on performance, our findings show that adopting stretch organizational goals leads to high performance for a few organizations while worsening performance for others. Here, the positive main effect of stretch goals is weak and largely overwhelmed by the impact of goal difficulty on performance variance. We conjecture that the more sophisticated the mental models of the managers relative to the complexity of the situation, the more forgiving the environment, the greater the resources and capabilities of the organization, the stronger the positive main effect of stretch goals on performance will be. It is perhaps not surprising that high-performing organizations, including, for example, NASA during the 1960s, or General Electric and Wal-Mart more recently, respond positively to stretch goals. However, for organizations in dynamically complex environments, where learning is difficult, actions are expensive and largely irreversible, and organizational capabilities and slack are poor, the effect of stretch objectives is likely to increase performance variability without much impact on mean performance. Thus if the organization's performance is poor relative to the industry mean, setting stretch goals is unlikely to lead to high performance. Perhaps stretch goals are a rule for "the rich getting richer". Exploring these possibilities is likely to be a fruitful area for future research.

In addition, there is increasing concern that stretch goals ultimately lead management to adopt actions and strategies that damage organizations' long-term health (Fuller and Jensen 2010, Ordóñez et al. 2009, Rahmandad 2012). For example, the President of Toyota Motor Corporation recently blamed the company's stretch goals for rapid growth and market share, for slipping safety standards, and publicly apologized for accidents caused by problems with Toyota vehicles that led to millions of recalled vehicles and damaged Toyota's brand image and sales (Kubo and Crawley 2010).

In general, the results show that stretch goals lead to a major increase in performance variance, with at best a limited increase in the performance for the median firm. Greater performance variance without an increase in expected outcomes amounts to an increase in risk and is not a choice most managers would make. Some managers might find the increase in risk attractive and adopt stretch

goals if they expect to be one of the market leaders in the high performing tail of the distribution, that is, if they discount the possibility of failure. Research shows most people are unrealistically optimistic about their position in a distribution of peers on almost any positive trait or ability (e.g., Svenson 1981, Weinstein 1980). Alternatively, as suggested by prospect theory, the higher risks associated with aggressive stretch objectives might be adopted even if there is no positive main effect if organizations perceive themselves to be in the domain of losses, for example, if senior leadership perceives that they must improve dramatically or face failure, or if senior managers perceive that without dramatic improvements they will lose the potential value of their stock or options. Experiments exploring the potential interactions of incentives and perceived self-efficacy with stretch goals would provide insight into these possibilities.

Overall, our findings show that stretch goals are not a 'rule for riches' and that organizations should carefully assess the complexity of their environment, the quality of their mental models, and their robustness to failed strategies before adopting stretch goals. Setting stretch goals implies that managers would frequently fail to meet their targets. Only thirteen percent of organizations met the long-run stretch goal in the experiment. How would an organization retain and motivate its managers who continuously fail to meet their goals? Do stretch goals increase turnover amongst senior managers, resulting in churn and loss of organizational memory?

The increasing use of stretch goals within organizations intensifies the need for empirical research on the effects of stretch goals to guide managerial actions. Stretch goals yield positive benefits for some organizations, but do not benefit all. Future research could build on our findings by exploring in more detail the balance between setting stretch versus moderate goal levels and identifying the conditions under which stretch goals are most appropriate and beneficial.

References

Andrews, K. R. 1971. *The concept of corporate strategy*. Dow Jones - Irwin, Homewood, IL.

Audia, P. G., E. A. Locke, K. G. Smith. 2000. The paradox of success: An archival and a laboratory study of strategic persistence following radical environmental change. *The Academy of Management Journal*. **43**(5) 837-853.

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.

Bartlett, C. A. 2000. *Ge Compilation: Jack Welsh, 1981-99*. Harvard Busn School.

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. In M. Hitt, Freeman, R., Harrison, J. (Eds.), *The Blackwell Handbook of Strategic Management*: 259-288. Malden, MA: Blackwell Publishers.

Button, K. 2012. Low-Cost Airlines. *Transportation Journal*. **51**(2) 197-219.

Chattopadhyay, P., W. H. Glick, C. C. Miller, G. P. Huber. 1999. Determinants of executive beliefs: Comparing functional conditioning and social influence. *Strategic Management Journal*. **20**(8) 763-789.

Chesney, A. A., E. A. Locke. 1991. Relationships among goal difficulty, business strategies, and performance on a complex management simulation task. *Academy of Management Journal*. **34**(2) 400-424.

Cohen, M. D. 1984. Conflict and complexity: Goal diversity and organizational search effectiveness. *The American Political Science Review* 435-451.

Collins, J. C., J. I. Porras. 2002. *Built to last*. Collins Business, Fulham UK.

Conover, W. J., M. E. Johnson, M. M. Johnson. 1981. A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics* 351-361.

Cronin, M. A., C. Gonzalez, J. D. Sterman. 2009. Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organizational Behavior and Human Decision Processes*. **108**(1) 116-130.

Cyert, R., J. March. 1963. *A Behavioral Theory of the Firm*. Blackwell Publishers Inc., Cambridge, MA.

Denrell, J. 2008. Organizational risk taking: adaptation versus variable risk preferences. *Industrial and Corporate Change*.

Diehl, E., J. D. Sterman. 1995. Effects of Feedback Complexity on Dynamic Decision Making. *Organizational Behavior and Human Decision Processes*. **62**(2) 198-215.

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.

Erez, M., I. Zidon. 1984. Effect of goal acceptance on the relationship of goal difficulty to performance. *Journal of Applied Psychology*. **69**(1) 69-78.

Fang, C. 2012. Organizational Learning as Credit Assignment: A Model and Two Experiments. *Organization Science*. **23**(6) 1717–1732.

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.

Gary, M. S., R. E. Wood. 2011. Mental models, decision rules, and performance heterogeneity. *Strategic Management Journal*. **32**(6) 569-594

Graham, A., J. Morecroft, P. Senge, J. Sterman. 1992. Model-supported case studies for management education. *European Journal of Operational Research*. **59**(1) 151-166.

Grant, R. M. 2010. *Contemporary Strategy Analysis*. Jogn Wiley & Sons Ltd., Chichester, U.K.

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.

Hamel, G., C. K. Prahalad. 1993. Strategy as stretch and leverage. *Harvard Business Review*. **71** 75-84.

Hofer, C. W., D. E. Schendel. 1978. *Strategy formulation analytical concepts*. West Publishing Co, St Paul Minn.

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.

Janis, I. L., L. Mann. 1977. *Decision making: A psychological analysis of conflict, choice, and commitment*. Free Press, New York, NY.

Jordan, A., P. Audia. 2012. Self-enhancement and learning from performance feedback. *Academy of Management Review*. **37**(2).

Kahneman, D., A. Tversky. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*. **47**(2) 263-292.

Klein, H. J., J. S. Kim. 1998. A field study of the influence of situational constraints, leader-member exchange, and goal commitment on performance. *Academy of Management Journal* 88-95.

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.

Kubo, N., J. Crawley. 2010. *Toyota says growth trumped safety, vows changes*.

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., D. B. Montgomery. 1987. Learning from strategic success and failure. *Journal of Business Research*. **15**(6) 503-517.

Larrick, R., C. Heath, G. Wu. 2009. Goal-induced risk taking in negotiation and decision making. *Social Cognition*. **27**(3) 342-364.

Levinthal, D. A., J. G. March. 1993. The myopia of learning. *Strategic Management Journal*. **14**(S2) 95-112.

Levitt, B., J. G. March. 1988. Organizational learning. *Annual Review of Sociology*. **14** 319-340.

Locke, E., G. Latham. 1990. *A theory of goal setting and task performance*. Prentice-Hall, Englewood Cliffs, NJ.

Locke, E., G. Latham. 2002. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*. **57**(9) 705-717.

Locke, E., G. Latham, M. Erez. 1988. The determinants of goal commitment. *Academy of Management Review*. **13**(1) 23-39.

Locke, E. A. 1982. Relation of goal level to performance with a short work period and multiple goal levels. *Journal of Applied Psychology*. **67**(4) 512-514.

Locke, E. A., K. N. Shaw, L. M. Saari, G. P. Latham. 1981. Goal setting and task performance: 1968-1980. *Psychological Bulletin*. **90**(1) 125-152.

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.

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.

Miller, K., W. Chen. 2004. Variable organizational risk preferences: Tests of the March-Shapira model. *The Academy of Management Journal*. **47**(1) 105-115.

Mishina, Y., B. Dykes, E. Block, T. Pollock. 2010. Why "Good" Firms Do Bad Things: The Effects of High Aspirations, High Expectations and Prominence on the Incidence of Corporate Illegality. *Academy of Management Journal*. **53**(4) 701–722.

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. *The Academy of Management Perspectives ARCHIVE*. **23**(1) 6-16.

Oxnard, T. 2004. Stretch!: How Toyota reaches for big goals. *Supply Chain Management Review*. **8**(2) 28-35.

Paich, M., J. D. Sterman. 1993. Boom, Bust, and failures to learn in experimental markets. *Management Science*. **39**(12) 1439-1458.

Peters, T., R. Waterman. 1982. *In Search of Excellence: Lessons from America's Best-Run Corporations*. Harper and Row, New York.

Quindlen, R., H. Greenberg. 2000. *Confessions of a venture capitalist: inside the high-stakes world of start-up financing*. Warner Books.

Rahmandad, H. 2012. Impact of Growth Opportunities and Competition on Firm-Level Capability Development Trade-offs. *Organization Science*. **23**(1) 138-154.

Rahmandad, H., N. Repenning, J. D. Sterman. 2009. Effects of Feedback Delay on Learning. *System Dynamics Review*. **25**(4) 309-338.

Rousseau, D. M. 1997. Organizational behavior in the new organizational era. *Annual Review of Psychology*. **48** 515-546.

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. *The Academy of Management Journal ARCHIVE*. **49**(4) 779-795.

Simon, H. 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.

Slater, R. 1999. *Jack Welch and the GE way: Management insights and leadership secrets of the legendary CEO*. McGraw-Hill Companies, New York.

Smith, K. G., E. A. Locke, D. Barry. 1990. Goal setting, planning, and organizational performance: An experimental simulation. *Organizational Behavior and Human Decision Processes*. **46**(1) 118-134.

Sterman, J. D. 1988. *People Express Management Flight Simulator* currently available from Strategy Dynamics http://www.strategydynamics.com/microworlds/people-express/, (software originally developed in 1988, Cambridge MA).

Sterman, J. D. 1989. Misperceptions of Feedback in Dynamic Decision making. *Organizational Behavior and Human Decision Processes*. **43**(3) 301-335.

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. 2010. Does formal system dynamics training improve people's understanding of accumulation? *System Dynamics Review*. **26**(4) 316-334.

Svenson, O. 1981. Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*. **47**(2) 143-148.

Thompson, K. R., W. A. Hochwarter, N. J. Mathys. 1997. Stretch targets: What makes them effective? *Academy of Management Executive*. **11**(3) 48-60.

Weinstein, N. D. 1980. Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*. **39**(5) 806-820.

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.

Wood, R. E., E. A. Locke. 1990. Goal setting and strategy effects on complex tasks. In B. Staw, Cummings, L. L. (Eds.), *Research in Organizational Behavior*, Vol. 12: 73-109.
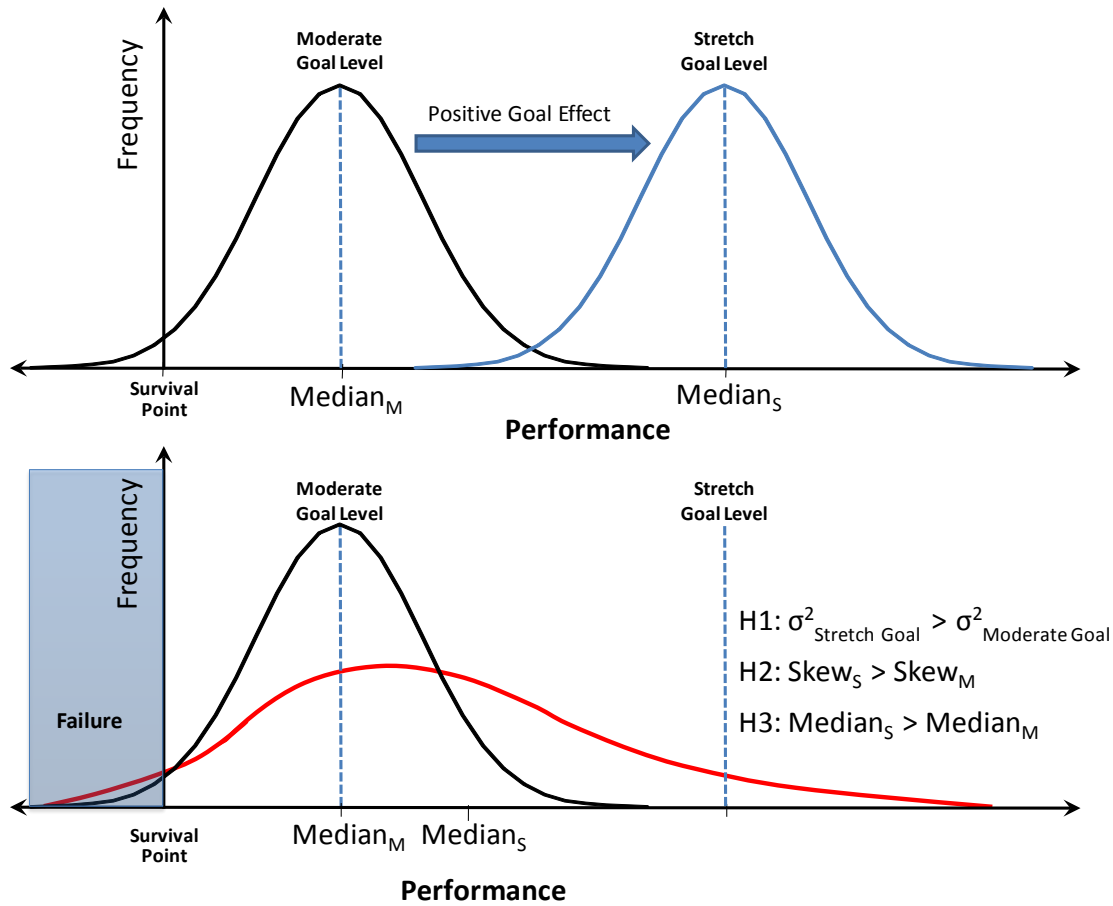
Figure 1 Illustration of hypothesized effects of stretch goals on performance level, variance, and skewness.

In the figure: H1: $\sigma^2_{Stretch\ Goal} > \sigma^2_{Moderate\ Goal}$, H2: $Skew_S > Skew_M$, H3: $Median_S > Median_M$.
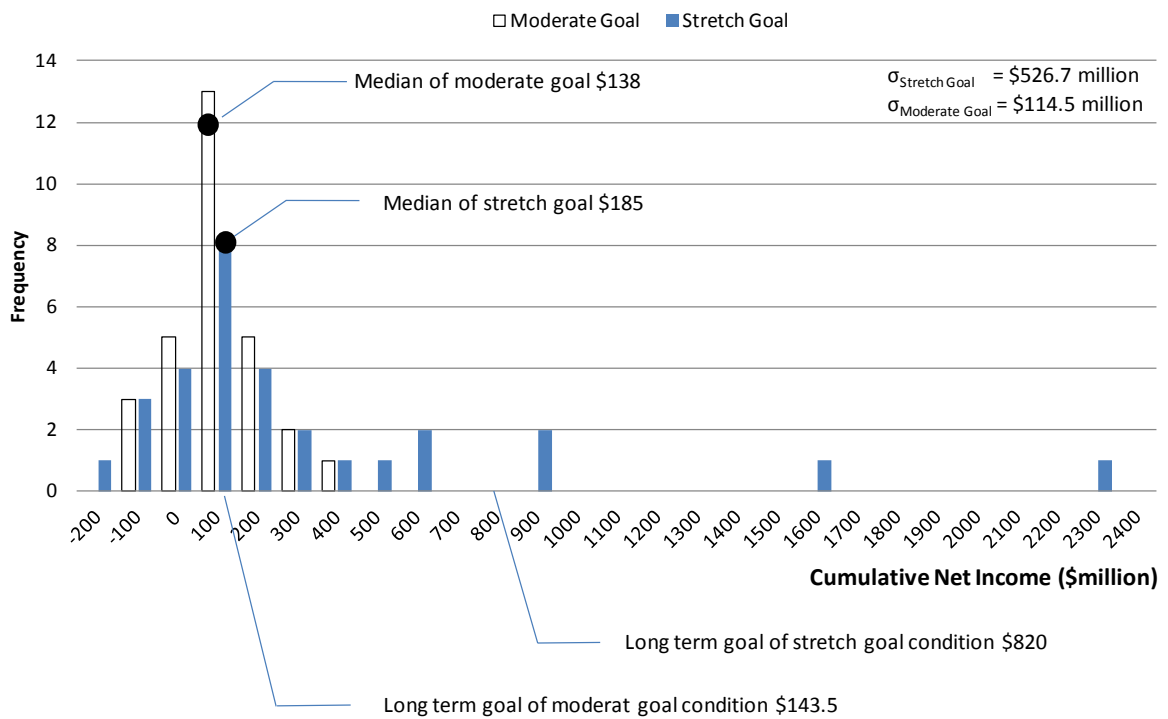
Figure 2 Performance distribution at the end of Year 10 for Stretch and Moderate Goal Conditions

In the figure: $\sigma_{Stretch\ Goal} = \$526.7$ million, $\sigma_{Moderate\ Goal} = \$114.5$ million. Median of moderate goal $138. Median of stretch goal $185. Long term goal of stretch goal condition $820. Long term goal of moderat goal condition $143.5.
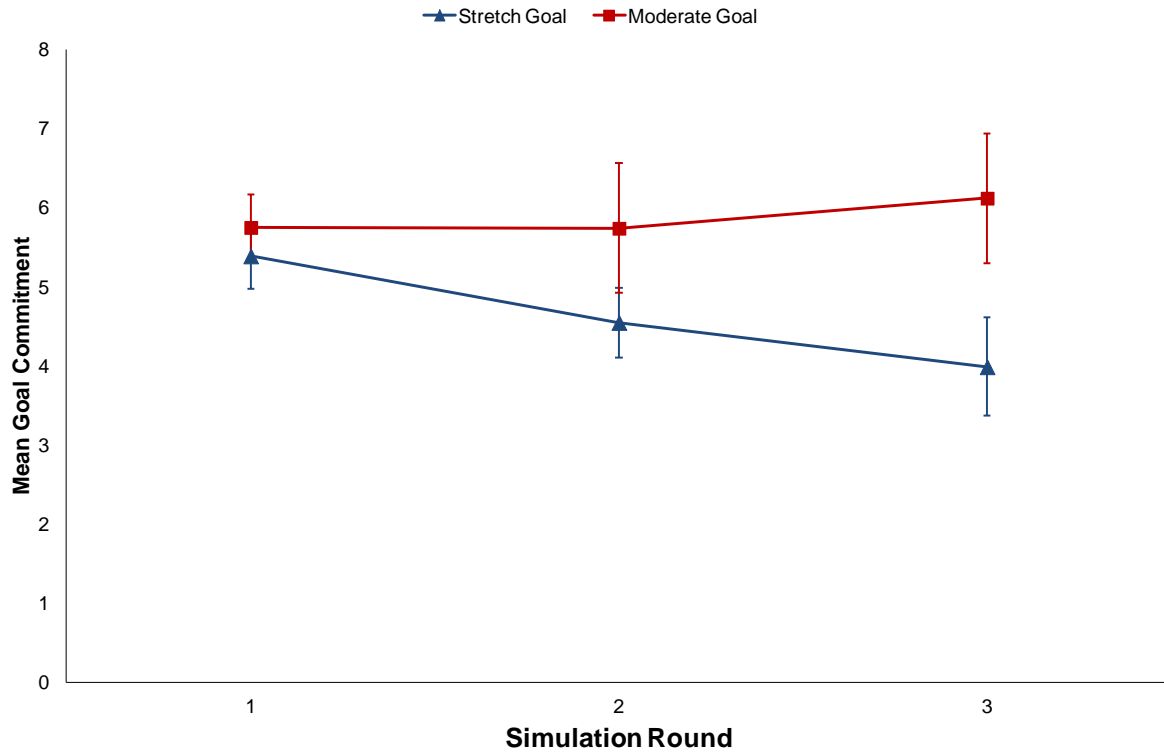
Figure 3 Mean Goal Commitment (with error bars of +/- two standard errors) across the three simulation rounds for stretch and moderate goal conditions. No significant differences between goal conditions for the first two simulation rounds, but significantly different for the third simulation round ($F_{[1, 57]} = 17.49$, $p < 0.001$).

Table 1 Payoffs for assigned moderate goal levels over the 10 year simulation

By the end of	Cumulative Net Income Goal ($million)	Your Actual Cumulative Net Income ($million)	Payment for achieving target
Year 1 Qtr 4	No annual goal No Payment		
Year 2 Qtr 4			
Year 3 Qtr 4	18.8		$2.00
Year 4 Qtr 4	27.7		$2.00
Year 5 Qtr 4	39.2		$2.00
Year 6 Qtr 4	54.4		$2.00
Year 7 Qtr 4	72.0		$2.00
Year 8 Qtr 4	92.4		$2.00
Year 9 Qtr 4	116.0		$2.00
Year 10 Qtr 4	143.5		$6.00

Table 2 Payoffs for assigned stretch goal levels over the 10 year simulation

By the end of	Cumulative Net Income Goal ($million)	Your Actual Cumulative Net Income ($million)	Payment for achieving target
Year 1 Qtr 4	No annual goal No Payment		
Year 2 Qtr 4			
Year 3 Qtr 4	31.5		$2.00
Year 4 Qtr 4	56.6		$2.00
Year 5 Qtr 4	99.2		$2.00
Year 6 Qtr 4	171.6		$2.00
Year 7 Qtr 4	269.3		$2.00
Year 8 Qtr 4	401.3		$2.00
Year 9 Qtr 4	579.5		$2.00
Year 10 Qtr 4	820.0		$6.00

Table 3 Descriptive statistics for performance in stretch and moderate goals, and test statistics for Fligner-Killeen and Levene tests.

Year	Assigned Goal		Median Performance		Std. Dev.		Fligner-Killeen Test[a]	Levene Test[a]
	Moderate	**Stretch**	**Moderate**	**Stretch**	**Moderate**	**Stretch**		
0	No Goal	No Goal	0	0	0.00	0.00		
1	No Goal	No Goal	3	2	2.7	6.4		
2	No Goal	No Goal	13	10	8.2	24.3		
3	18.8	31.5	26	22	13.0	31.2	3.52	4.75
4	27.7	56.6	34	36	19.3	36.5	1.98	3.10
5	39.2	99.2	52	61	27.6	47.8	3.50	3.75
6	54.4	171.6	65	78.5	38.7	84.8	6.82*	7.63*
7	72.0	269.3	82	105	52.2	139.0	7.63*	10.59*
8	92.4	401.3	99	133	69.9	231.0	9.07*	12.33**
9	116.0	579.5	117	160	91.3	371.2	8.47*	12.67**
10	143.5	820.0	138	185	114.5	526.7	8.32*	13.15**

Sample size for each year equals 29 in moderate goal condition and 30 in stretch goal condition.
a) p-values corrected for multiple dependent hypotheses testing using the Holm-Bonferroni adjustment
*$p < 0.05$; **$p < 0.01$.

Table 4 Number of organizations in each performance segment, with segments defined under moderate condition

	Above survival and below moderate goal	Meet or exceed moderate goals	TOTAL
Moderate Goal Condition	12	12	24
Stretch Goal Condition	5	18	23
TOTAL	17	30	47