The effects of a goal setting intervention on productivity and persistence in an analogue work task

Tammemagi, Triona; O'Hora, Denis

2013-02-27

Tammemagi, T, O'Hora, D, Maglieri, KA (2013) 'The effects of a goal setting intervention on productivity and persistence in an analogue work task'. Journal Of Organizational Behavior Management, 33:31-54.

Taylor & Francis

http://dx.doi.org/10.1080/01608061.2013.758001

http://hdl.handle.net/10379/5823

http://dx.doi.org/10.1080/01608061.2013.758001
The Effects of a Goal Setting Intervention on Productivity and Persistence in an Analogue Work Task

Triona Tammemagi and Denis O’Hora
National University of Ireland, Galway
Kristen A. Maglieri
University of Dublin, Trinity College

Author Note

Triona Tammemagi and Denis O’Hora, School of Psychology, National University of Ireland, Galway; Kristen Maglieri, School of Psychology, University of Dublin, Trinity College.

This investigation was supported in part by the Chris Anderson Research award grant, awarded by the Organizational Behavior Management Network, and the Irish Research Council for the Humanities and Social Sciences.

Correspondence concerning this article should be addressed to Triona Tammemagi, School of Psychology, National University of Ireland Galway, Galway, Ireland. E-mail:

t.tammemagi1@nuigalway.ie
Abstract

The study sought to quantify the beneficial effect of goal setting on work performance, and to characterize the persistence or deterioration of goal-directed behavior over time. Twenty-six participants completed a computer-based data entry task. Performance was measured during an initial baseline, a goal setting intervention that consisted of either a high, unattainable goal (high goal condition) or a low, attainable goal (low goal condition), followed by a return to baseline, and a second goal setting intervention (the alternate goal to the first goal). In the fifth condition, each participant was given the choice to work in either the high or low goal condition. Greater performance increases were reliably observed during the high goal condition than during the low goal condition, but patterns of persistence or deterioration varied across participants. The implications of the findings for the development and understanding of goal setting interventions in the workplace are explored.

Keywords: goal setting, goal-directed behavior, persistence, productivity, performance
The Effects of a Goal Setting Intervention on Productivity and Persistence in an Analogue Work Task

Goal setting is one of the most widely used and effective interventions to increase work performance. There have been numerous social, cognitive and behavioral theories posited to account for goal setting and to explain goal-directed behavior (e.g., Austin & Vancouver, 1996; Latham, 2003; Latham & Locke, 1979; Steers & Porter, 1974; Waldersee & Luthans, 2001). Industrial-organizational researchers have focused, to a large extent, on the effects of goal level (Locke & Latham, 2002), and goal specificity (Latham, 2003) on performance. Results have consistently shown that high specific goals tend to lead to increased performance (Jackson & Zedeck, 1982; Locke & Latham, 2002; Steers & Porter, 1974).

Behavioral accounts

Within the field of behavioral theory, a number of accounts have been provided to describe goal setting and goal-directed behavior. Early accounts described goals as both discriminative stimuli and conditioned reinforcers (Fellner & Sulzer-Azaroff, 1984; Huber, 1985). Fellner and Sulzer-Azaroff (1984) conducted a review of goal setting research from a behavior analytic perspective. They suggested that the combination of a goal and feedback places behavior under stimulus control, with the goal acting as the antecedent stimulus and feedback operating as reinforcement. According to this account, if meeting a goal is paired with a reinforcing consequence, then goal achievement itself can assume the properties of a conditioned reinforcing stimulus. Thus, according to Fellner and Sulzer-Azaroff, goals can function as both discriminative stimuli and conditioned reinforcers.

In contrast, some accounts of goal setting (e.g. Agnew & Redmon, 1993; Malott, 1993; Malott, Shimamune & Malott, 1993), have described goals as rules that exert control over behavior. The statement of a rule (a goal) alters the value of a delayed consequence by
making its influence immediate. Setting up a rule creates an immediate controlling circumstance that affects behavior until the delayed consequences can exert more direct control (Malott, 1993). Malott points out that the delay between the behavior and the outcome is crucial for direct contingencies of reinforcement, but has little effect when rules describe those contingencies – a goal is provided, and a week later the behavior occurs that will result in reinforcement. Given sizable and probable outcomes, people will follow rules even though those outcomes are delayed.

In Agnew’s (1997) account of the motivating operation in organizations, a goal is seen as a motivating operation (MO) that alters the value of a reinforcer, rather than signaling the availability of the reinforcer. To explain this, Agnew draws a distinction between discriminative stimulus accounts and MO accounts of goal-directed behavior. In a traditional discriminative stimulus account of the effect of feedback (as a reinforcer) on goal-directed behavior, feedback is differentially available in the presence of the goal, and the goal signals the availability of the reinforcer (feedback). However, in a typical work setting, feedback is present before the goal setting intervention has been implemented, but is more valuable when the goal has been put in place. Agnew explained that discriminative stimulus accounts can account for some goal setting effects when the reinforcer is not available in the absence of the goal, but viewing a goal as a motivating operation can better explain more instances of goal-directed behavior.

O’Hora and Maglieri (2006) provided an account of goal setting that focused on the content of goal statements and the organizational contingencies that maintain goal-directed behavior. This account was based on relational frame theory, a behavioral theory that describes the core processes of language (e.g., semantics) in terms of relations between words and stimuli in the world. By examining the relations established by goal statements, their intention was to provide a functional explanation of how goal statements allow reinforcement
of goal-directed behavior. O’Hora and Maglieri suggested that goal statements work by setting up comparison relations between the goal statement and ongoing self-statements that describe current levels of performance. They described an example in which a goal statement is given to employees: “For this week, everyone is expected to make 60 sales”. As the week progresses and the employee works, the employee issues self-statements (e.g. “I have made 20 sales”), and compares current performance to goal level performance. The comparison relations between the current performance and the goal decrease gradually as the person works. For example, if the employee states “I have made 30 sales”, the ‘less than’ comparison between the performance and the goal (60 sales) has decreased to 30. Later, after further work, the employee can state “I have made 40 sales”, and the comparison relation has decreased further to 20. As a person works, the self-statements establish smaller and smaller ‘less than’ relations between the current performance and the goal. O’Hora and Maglieri have suggested that reducing these ‘less than’ relations serves as derived reinforcement of goal-directed behavior.

**Productivity and Goal Level**

Locke and Latham are two of the foremost theorists to contribute to the development of goal setting research (e.g., Latham, 2003; Latham & Locke, 1979; Locke, 1967; Locke & Latham, 2002). Latham and Locke (1979) proposed that there is a positive linear relationship between goal level and performance. As goals increase, so does task performance. Latham (2003) reported on a group of loggers who set a specific high goal around a number of trees to cut down in a week. The results were an increase in both attendance and performance. Similarly, more specifically defined goals produce a greater level of performance than loose “do your best” goals (Locke & Latham, 2002). Locke and Bryan (1967) conducted a series of studies in which participants were asked to complete addition, perceptual speed and psychomotor coordination tasks. In the first of these experiments, participants were given a
number task and were assigned either a specific hard goal or a “do your best” goal. Participants given specific hard goals performed at a higher level in the first 15 minute period, and this size difference increased throughout the 90 minute study. Participants given the specific hard goal scored an average of 9% higher than the group that was told to do their best. Latham and Baldes (1975) conducted a study implementing specific hard goals with truck drivers to improve performance. Previous data had revealed that the net weight of logging trucks had frequently fallen short of the legal maximum weight, and the aim of the goal setting intervention was to increase the net weight of trucks transporting logs. The authors introduced a specific hard, yet attainable goal. Prior to implementation the truckers had been told to “do their best” and the percent net weight of 36 logging trucks was reported as between 55 and 65. Results showed a substantial increase in performance, with a rise in net weight to approximately 90%, which in this case held across eight months.

The first aim of the current study was to replicate previous findings that high specific goals lead to greater performance (Jackson & Zedeck, 1982; Locke & Latham, 2002; Steers & Porter, 1974). The aim was to test whether the introduction of a goal would significantly affect productivity. Productivity was measured as mean correct responses per condition. It was expected that performance would significantly increase in the presence of the high goal, and if mean baseline performance was below the low goal level, performance would increase when the low goal was introduced.

**Persistence and Goal Difficulty/ Attainability**

Goal setting theory suggests that difficult goals lead to higher performance levels; however exactly how difficult a goal should be is unclear (Locke & Latham, 1990). Researchers have shown that goals that are too high may not encourage long-term increase in performance. See, Heath, and Fox (2006) conducted a study testing persistence of goal-directed behavior utilizing two unattainable goals, one that was ‘just out of reach’ and the
other that was a high unattainable goal. Participants were required to “sit” on air with their backs straight against a wall and knees bent to a 90 degree angle, either for 6 minutes or 4.5 minutes. Participants in the low goal condition persisted with the task longer than those in the high goal condition (when tests of median performance were analyzed); however when individual data was examined the authors noted that individual differences existed in patterns of persistence. Low performers persisted less in the high goal condition, and high performers persisted more with the high goal. The authors asserted that a goal that is “well beyond reach of most people” may result in a drop in performance of those people (p.25).

Negotiation research has contributed significantly to the study of abandoning goals (i.e. failure to persist), or ‘resistance point’ (e.g., Zetik & Stuhlmacher, 2002). In line with See et al.’s (2006) study, Wrosch, Scheier, Miller, Schulz, and Carver (2003) suggested that one factor leading to the abandonment of a goal may be the individual’s perception of the probability of goal attainment. To put it simply, a person who believes that a goal is attainable is less likely to abandon it, and more likely to persist. Hatzigeorgiadis (2006) examined the effects of unattainable goals on a group of sports students in a rowing task. The researcher split participants into an ‘attainable goal’ group and an ‘unattainable goal’ group. The study primarily measured approach and avoidance coping as a function of goal attainment expectancies; however a secondary level of analysis examined the rowing performance of the two groups. In the attainable goal group, where scores were higher for effort and planning, the rowing tempo significantly increased in the last leg of their task. Conversely in the unattainable group where scores were higher on behavioral disengagement, rowing tempo did not significantly change throughout the task. When the high unattainable goal was in effect, performance remained the same throughout the task. A high goal may seem attainable over a short period of time, but over a longer period of time, as the person’s
performance does not significantly progress towards the goal, performance may stabilize or decrease.

The second aim of the current study was to test for persistence within a high, unattainable goal condition. It was proposed that at some point, in the absence of goal attainment, goal-directed behavior would extinguish, and the study aimed to pinpoint when this would occur. According to behavioral accounts of goal setting, behavior occasioned by verbal rules will extinguish over time in the absence of reinforcement (Fellner & Sulzer-Azaroff, 1984). O’Hora and Maglieri (2006) have proposed that achieving goal levels periodically is necessary in order to maintain performance. Unattainable goals may give rise to high levels of performance, but the level of behavior should decrease because the derived reinforcing function of the employee’s self-statements may extinguish over time. For the purpose of the current study, persistence was defined and measured as a stable or positive trend across sessions within a goal condition. It was expected that performance would drop within the high goal condition, which would be determined by a negative persistence score.

The study employed a reversal design, to investigate whether performance would revert to baseline levels when the goal was no longer presented. The study also examined whether participants chose a low, easily attainable goal, or the alternative, a high, yet unattainable goal. A low goal allows easier access to reinforcement (goal attainment and/or feedback); therefore it was predicted that participants would choose a low goal condition when presented with a choice.

Method

Participants

Twenty-six participants, ranging in age from 18 to 22 years (8 males and 18 females), took part in the study. Participants were first year undergraduate psychology students at a
public university in Ireland, recruited through an online university system. Participants earned course credit for participating in the study.

**Apparatus and Setting**

The experimental task was designed in Visual Basic 6.0 to mimic a medical data entry task, and participants used both the keyboard and mouse to respond. The program presented all stimuli and recorded all responses. The task was presented on a Fujitsu Siemens personal computer with a 17inch CRT screen, in a laboratory room in the School of Psychology at a large university in Ireland, for a total duration of 1 hour 20 minutes. The experimenter remained in the room for the training component of the study, and the participant was alone for the testing phase.

**Design**

**Dependent variables.** Task performance was measured in two ways: (a) Productivity, which was defined as the mean number of correct responses per condition, and (b) Persistence, defined as the change in number correct per session across sessions within a condition. Trend was measured across sessions within a condition, and a stable or positive trend indicated persistence.

**Independent variables.** The independent variable goal was presented in three levels: (a) no goal, (b) low (attainable) goal, and (c) high (unattainable) goal. Both the low and high goals were selected from the results of pilot studies.

**Research design.** The experiment tested the effects of goals on task performance over time. A single subject modified reversal design (ABACX) was used, which consisted of a standard ABAC reversal design followed by a final choice condition. Conditions proceeded as follows:

- **a) Baseline,** in which the participant was not provided with any goal;
b) First goal, in which a manager (either Bob or Todd) provided a high or low goal;

c) Second baseline;

d) Second goal, in which the alternative goal condition to b) was presented; and

e) Choice, in which the participant was presented with a choice between the
managers from the low or high goal conditions.

The presentations of the low and high goal conditions were counterbalanced across
participants to avoid order effects.

Procedure

Experimental task. A data entry task was utilized, which was designed to simulate
typical electrocardiogram (ECG) data a nurse might enter into a database system. The
simulation was modified from a task used in a prior experiment by Maglieri (2007). The
screen contained fictional medical information related to an ECG reading, and participants
were required to examine the data provided for each patient and decide whether it was within
range or out of range for each patient. Participants followed a series of steps to enter the
correct data. Figure 1 shows the work task screen.

- - - - - - - - - - - - - - - - - - - - - - - - - - -
Insert Figure 1 about here
- - - - - - - - - - - - - - - - - - - - - - - - - - -

The participant entered the patient’s ID number (‘1’ in Figure 1) in the PATIENT ID
box (2) in the center of the screen. The participant then checked the patient’s gender (3), and
his or her QT interval number (a measure of time between the Q wave and the T wave in the
heart’s electrical activity), (4) and compared it to the relevant range of numbers in the box (5)
in the center of the screen (i.e. for female patients, compare the FEMALE range, for males,
the MALE range). The participant ticked the appropriate dot (either within or out of range)
to classify the data (6). If the patient’s QT interval was within the range for his or her gender, then the participant ticked the dot next to ‘within range’. If not, the participant ticked ‘out of range’. When the above steps were completed the participant clicked on SUBMIT in the center of the screen (7). If the patient’s data was entered correctly the number at the bottom of the screen increased by one (8). The screen was then refreshed and the participant began the next trial. Each work session lasted for four minutes, with three work sessions per condition. At the end of each work session, feedback was presented to the participant indicating how many patients were entered correctly during that work session, and how many were entered incorrectly. Figure 2 displays the screen presented to participants at the end of each condition. The screen presented feedback on amount correct per work period, overall amount correct for that condition, and the number of errors for that condition.

Training. During the training session, a pop-up window appeared on screen and provided the participants with detailed instructions about the task (see Appendix for instructions). Participants were encouraged to stop at any time to ask questions. The training session consisted of two 1-minute conditions. The first condition was a reduced-length version of the baseline condition. The second condition was a reduced-length version of the last ‘choice’ condition, in which participants were asked to choose between either a high or low goal. These conditions will be detailed fully below. The experimenter ensured that the participant could engage in the required responses before entering the testing phase. When the training was complete the experimenter set up the testing program and informed the participant that the experimenter would be leaving the room.
**Testing.** Testing lasted a total of 60 minutes, with five 12-minute conditions. Each condition consisted of three 4-minute work sessions. Pilot studies indicated that four minute sessions provided sufficient variability in performance across participants, and three presentations of each session was short enough to complete within scheduled timeslots.

**Baseline.** Participants classified the data as stated previously, and could see the number of patients’ data they were entering correctly throughout the entire work session (number (8) in Figure 1). Feedback on the number of correct and incorrect entries was also displayed following each work session. At the end of the baseline condition, feedback was given on the number of correct and incorrect entries per work session, as well as the total correct for the entire condition.

**Low goal.** When presented with the low goal condition, participants were given instructions from a virtual manager, Bob, to enter 20 patients’ data correctly for that work session. The instruction was given via a pop-up window on the screen, and the instruction remained on the task screen throughout the condition. As in baseline, feedback was given on the screen throughout the task, informing participants on correctly entered patient data. Feedback on the number of correct and incorrect responses was also given at the end of each work session. At the end of the condition, feedback was given on the number of correct and incorrect responses per work session, along with a cumulated correct score for all conditions.

**Baseline.** This condition was identical to the first baseline condition. No goal was given; however as in previous conditions, the participant received feedback during the work session, at the end of the work session, and at the end of the condition.

**High goal.** During the high goal condition, participants were given instructions from a different virtual manager named Todd, who asked for 40 correct data entries. Again, each participant received feedback on correct entries throughout, and feedback on the number of correct and incorrect data entries following each work session. The manager’s instructions
remained on the task screen for the duration of the condition. At the end of the condition, participants received feedback on the number of errors and correct responses for each work session, as well as a cumulated number of correct responses for the entire condition.

**Choice.** The fifth condition was a choice condition in which participants were required to choose between managers Bob (low goal) and Todd (high goal). Whichever manager they selected dictated which goal condition they would be exposed to. The manager’s instructions remained on the screen throughout the task, along with feedback on the number of correct responses. At the end of the work session participants received feedback on the number of correct and incorrect data entries for that work session. At the end of the condition, feedback was presented on the number of errors and correct responses per session, along with a cumulated correct score for the entire experiment.

Participants had the option to take a break between sessions or between conditions. The experimenter remained outside the room for the duration of the testing. At the conclusion of the study, the experimenter debriefed the participant on the goal of the study, and thanked the participant for taking part.

**Results**

It was expected that performance would significantly increase with the introduction of a high performance goal. A 2 (Group) x 4 (Phase) mixed ANOVA was conducted to examine whether performance differed between the two groups (high goal first and low goal first), and across the first four phases. There was no significant difference in overall performance between the two groups, $F(1,24) = 0.14, p = 0.71, \eta_p^2 = 0.01$. A main effect of Phase was observed, $F(3,72) = 33.25, p < 0.001, \eta_p^2 = 0.58$, indicating that performance was significantly different across phases. Mean performance was highest in the fourth phase (second goal condition; $M = 25.19, SD = 5.9$) and lowest in the first phase (baseline; $M = 21.12, SD = 4.29$). A significant Group x Phase interaction was also observed, $F(3,72) =$
3.29, \( p = 0.03, \eta^2_p = 0.12 \). Figure 3 displays the average performance across phases (including the fifth ‘choice’ phase) by the high goal first and low goal first groups. The general increase in performance demonstrated the expected effect of the presentation of the high goal. During Phase 2, performance increased to a greater extent for the high goal first group than for the low goal first group. For the low goal first group, performance increased in Phase 2, and increased again when participants were presented with the high goal in Phase 4.

To further analyze the interaction effect reported above, 2 one-way (Phase) within-subjects ANOVAs were conducted to investigate how the introduction of a high or low goal condition affected performance for the high goal first and the low goal first group.

**High goal first.** When the high goal was introduced first, a significant difference in performance across phases was observed, \( F(3,36) = 28.18, \ p < 0.001, \eta^2_p = 0.70 \). Performance was highest in Phase 3 (second baseline condition; \( M = 25.38, SD = 4.60 \)) and lowest in Phase 1 (first baseline; \( M = 20.95, SD = 4.28 \)). A series of planned comparison \( t \)-tests were conducted to examine significant differences between conditions (see Table 1). Significant differences were observed between the first baseline condition and all later conditions, demonstrating that performance increased on introduction of the high goal, and then remained stable as shown in Figure 3. Control for multiple testing was ensured using Hochberg’s (1988) procedure for multiple tests of significance (critical \( p \)-value = 0.0125).
Low goal first. A further one-way (Phase) within-subjects ANOVA was conducted to examine the effect of the goal conditions on performance, when the low goal was introduced first. A significant difference in performance across phases was observed $F(3,36) = 11.77, p < 0.001, \eta^2_p = 0.50$. Performance was highest in Phase 4 (high goal condition; $M = 25.38, SD = 5.25$) and lowest in Phase 1 (first baseline; $M = 21.28, SD = 4.46$). Results from six planned comparison $t$-tests are presented in Table 2.

As with the previous group, control for multiple testing was ensured using Hochberg’s (1988) step-up sequential method (critical $p$-value = 0.0125). Significant differences were observed between the first baseline condition and all later conditions. Although the difference between the low and high goal was significant using a $t$-test, when we controlled for multiple comparisons this was not significant. In accordance with Hochberg’s correction method this also held true for differences between the second baseline and the high goal.

An increase in accurate responding may be achieved by reducing errors at the current rate, or by increasing the rate of correct responding. In order to establish how participant’s scores increased from baseline to the goal conditions, the relative accuracy across sessions was measured by calculating the percentage of correct responses per phase. Mean percent correct scores did not differ between phases for either the high goal first group or the low goal first group. Table 3 shows overall mean percent correct scores for each phase for the high goal first and the low goal first groups. In the high goal first group, only one participant’s mean percent correct score was below 80%. This participant’s mean percent correct scores were 100% in both the first baseline and the high goal conditions, 96% in the
second baseline, 88% in the low goal condition and 71% in the choice condition. Within the low goal first group, all participants’ mean scores were above 80% correct in each phase.

Insert Table 3 about here

Persistence Measure

In order to quantify trends within phases, slope values were calculated for each goal condition for each participant using least squares regression (see Tables 4 and 5 for slope values). These values were compared using two one-way within subjects ANOVAs, one for each group, to investigate whether trends varied significantly across phases. Upward or stable trends (positive or 0 slopes) indicated persistence. For the high goal first group a significant difference was found across phases, $F(3,36) = 4.82, p = 0.006, \eta_p^2 = 0.54$. A significant difference was found in trend between the first baseline condition (Phase 1) and the low goal condition (Phase 4), $p = 0.004$. The mean slope for the first baseline condition ($M = 1.46$) was strongly positive, and trend in the subsequent high goal condition ($M = 0.77$) was slightly positive. Mean trend was stable during the second baseline condition ($M = -0.08$), and then slightly negative in the low goal condition ($M = -0.31$). The initial strong positive trend was probably indicative of a practice effect during baseline. When the low goal was the first to be introduced, there was no significant difference found in trend across phases, $F(3,36) = 0.53, p = 0.66, \eta_p^2 = 0.12$.

Insert Table 4 about here

Insert Table 5 about here
Choice Condition

Twenty of the 26 participants (77%) chose the manager from the low goal condition in the final ‘choice’ condition. Five of the six participants that chose the manager from the high goal condition were from the low goal first group. Thus, these participants were exposed to the high goal condition as their second and most recent goal condition. For the low goal first group, mean scores did not differ between the second goal condition (high goal; $M = 25.38, SD = 5.25$) and the choice condition ($M = 24.72, SD = 5.00$). Similarly, for the high goal first group, mean scores did not differ between the second goal condition (low goal; $M = 25.00, SD = 5.00$) and the choice condition ($M = 26.28, SD = 5.52$). Interestingly, participants who chose the low goal condition continued to perform above the goal of 20 correct entries.

Single-subject Analysis

Due to the large number of participants, single subject graphs will not be presented for all participants. Analyses consisted of visual inspection of the data, along with inspection of slope values. To provide detail on individual performance, each participant’s phase mean, standard deviation and trend are provided in Table 4 for the high goal first group, and Table 5 for the low goal first group. Single-subject analysis further highlights the trends observed in the group data.

High goal first. For the high goal first group, 77% of participants demonstrated a positive trend in performance during Phase 1 (baseline). This may be due to an acquisition effect during initial introduction to the task. For 77% of participants, an increase in level was observed upon introduction of the high goal in Phase 2. A positive trend within the high goal condition was observed for 62% of participants. For 31% of participants, a negative trend in performance was observed during the high goal condition, while trend remained stable for
7% of participants (see Table 4). Only 15% of participants in Phase 3 (second baseline condition) showed a decrease in level of performance, whereas 39% of participants showed a negative trend within this condition. For the remaining participants, a stable (15%) or positive (46%) trend was observed. Figure 4 shows scores from one participant, with a negative trend in performance during the second baseline condition. With the introduction of Phase 4 (low goal), level of performance decreased for 23% of participants, and increased for 15% of participants. A positive trend was observed for 39% of participants, while a negative trend was found for 54% and a stable trend for 7% of participants (see Table 4).

---

Insert Figure 4 about here

---

**Low goal first.** Sixty-two percent of participants demonstrated a positive trend during Phase 1 (baseline). A negative trend was observed for 31%, while for 7% of participants, trend remained stable. Sixty-two percent of participants demonstrated an increase in level of performance with the introduction of Phase 2 (low goal condition). During Phase 2, a positive trend was observed for 38% of participants, while a negative trend was found for 31% of participants. The remaining 31% of participants demonstrated stable trend during the low goal condition. For the low goal first group, a return to baseline (Phase 3) resulted in a decrease in level of performance for 15% of participants, and a negative trend was observed for 54% of participants. The remaining participants demonstrated positive (39%) or stable (7%) trend within this phase (see Table 5). The introduction of Phase 4 (high goal) resulted in an increase in level of performance for 38% of the group. During Phase 4, 46% demonstrated a negative trend in performance (see Figure 5 for one of these participants); while 31% demonstrated a positive trend and trend remained stable for 23% of
participants. Figure 6 shows an example of a level increase for one participant with the introduction of the high goal in Phase 4.

Discussion

Overall findings demonstrated a greater increase in performance during the high goal condition than during the low goal condition; however patterns of persistence or deterioration in responding varied across participants. In the current study, three specific findings demonstrated the effects of high goals on performance. First, mean performance increased by 21% when a high goal was introduced after baseline (high goal first group), whereas mean performance only increased by 11% when a low goal was introduced (low goal first group). Second, for both groups, performance increased significantly upon introduction of the high goal (see Figure 3). Finally, in the low goal first group, performance was highest in the high goal condition, and in the high goal first group, performance increased in the high goal condition, and did not significantly increase beyond that level. These findings are broadly in line with previous research demonstrating that the provision of high specific goals increase performance compared to baseline levels (e.g., Latham & Locke, 1979; Locke & Latham, 2002; Steers & Porter, 1974). The current findings are in line with Locke and Bryan’s (1967) findings in which performance increased by 9% more when a specific difficult goal was introduced than when a “do your best” goal was introduced. In keeping with Locke and Bryan’s findings, performance increased immediately upon introduction of the high goal.
The within-phase trend was measured using slope values within the high unattainable goal condition in order to identify any decreases in persistence due to the unattainability of the high goal. It was expected that a negative trend would be observed across sessions in the high goal phase. According to O’Hora and Maglieri (2006), as participants were not getting closer to the goal, derived reinforcement would be undermined. Specifically, ‘less than’ relations between the goal statement and current self-statements of performance would not decrease sufficiently to serve as derived reinforcement. Although no participants reached the goal, a negative trend in performance was observed within the high goal condition for only 38% of participants. However, due to the constraints of the procedure, participants had relatively little exposure to the high goal condition (three 4-minute sessions). This may not have been sufficient contact with the condition for participants to determine that the high goal was unattainable, particularly when the high goal was the first to be introduced. When the high goal was presented as the second goal condition, a negative trend was observed for more participants (46%) than when it was presented as the first condition (31%). This may be evidence that with repeated exposure to the task, although there was an initial increase in productivity, the ‘less than’ relations between current self-statements and the goal may not have reduced significantly to increase performance further. Similar patterns of variation in persistence were obtained by See et al. (2006). Although the authors found that the median participant in the low unattainable group persisted for longer (113 seconds) with the task than the median participant in the high unattainable group (81 seconds), they noted that there was high variability between participants. These authors found that most participants persisted longer with a lower goal, but they noted that when mean performance was analyzed, the low performance in the higher goal condition was offset by the high performance of the top 15% of the group.
According to the O’Hora and Maglieri model, although high goals lead to higher levels of behavior, achieving goal levels periodically is necessary to maintain goal-directed behavior. The majority of behavioral accounts agree that goal attainment is necessary to maintain performance. Fellner and Sulzer-Azaroff (1984) have stated that some goals can function through rule-governed behavior, however, without the support of external contingencies, behavior under verbal rules may extinguish over time, as in the case of an unattainable goal. Unattainable goals may produce high levels of performance, but behavior will decrease over time if reinforcement for goal-directed behavior is not present. O’Hora and Maglieri propose that without such reinforcement the derived reinforcing function of the employee’s self-statements may extinguish over time. This effect was not reliably observed in the current study. The observed variability in persistence may have been due to participants’ differing behavioral histories with respect to goal-directed behavior, which gave rise to goal-directed behavior that was more or less resistant to extinction (lack of goal attainment). Short session length or insufficient exposure to the high goal condition might also have contributed to variability in persistence. Further studies might seek to expose participants to repeated work sessions over a longer period of time to further investigate the effect of an unattainable goal on persistence. Generally, in a work setting goals are in place for an extended period of time. It is difficult to generalize the current results to an organizational setting because of the brevity of the sessions; however the study provides material for planning future studies that would be better designed for answering the question of persistence with an unattainable goal.

In the final phase of the current study, 20 of 26 participants chose the low goal condition (involving the least effort to attain), when given the choice between a low or high goal condition. Locke (1967) suggested that choices should be given to determine preferred goal level. Studies have investigated the effect of choice when examining participation in
goal setting (Earley, 1985; Garland, 1983; Locke, Frederick, Lee & Bobko, 1984); however these studies have not examined a controlled forced choice, rather choosing an arbitrary goal level. Matching law (Herrnstein, 1961) states that the relative rate of reinforcement for a response will determine its strength or frequency. Specifically, when given a choice between two schedules of reinforcement, a participant will choose the schedule with the highest value (and least effort to attain). In the current study, however, no reinforcers were presented during baseline or goal phases. If goal attainment serves as a generalized or derived reinforcer (Fellner & Sulzer-Azaroff, 1984; O’Hora & Maglieri, 2006), then matching law predicts the observed performances. Interestingly, during the choice condition, five of the six participants that chose the high goal were presented with the high goal as their second goal condition, immediately before the choice condition was presented. This effect was unexpected, but future studies examining choice of goal statements might administer post-study questionnaires to shed light on this issue, should it recur.

The study employed a reversal design to investigate whether withdrawing the goal would cause performance to revert to baseline levels. In the low goal first group, a decrease in level from the goal condition to the second baseline was observed for only 15% of participants, but a negative trend in the second baseline condition was observed for 54% of participants. Within the high goal first group, there was a decrease in level between the goal condition and the second baseline for 15% of participants, and a negative trend in second baseline was observed for 39% of participants. The absence of a withdrawal effect may have been due to a lack of competing responses for the participant to make. In a typical work environment there are many stimuli present (e.g., internet, concurrent goals, unplanned interruptions), which may evoke competing responses that result in higher levels of reinforcement than achieving the task at hand. These stimuli may function as Motivating Operations, motivating escape behaviors that are incompatible with the targeted work
behavior (Olson, Laraway, & Austin, 2001). As these were not present in the current study, participants may have continued to work towards a higher level of data input, despite reaching the goal or even when the goal was removed. Future studies might focus on enhancing ecological validity, by providing alternative responses such as internet or phone access. It may also be beneficial to assess whether participants were self-setting goals during the experiment. Future studies may benefit from collecting self-report data following the task, investigating whether participants set their own goals during the study. Self-report data could also assess whether participants were attending to the instructions and feedback on-screen. If participants were not attending to the instructions or feedback on the screen, they may have continued to work towards a goal that was no longer in place or they may not have attended to the assigned goal level.

It was expected that when the low goal level was introduced first, those that performed below 20 during baseline would increase their performance to 20 to meet the goal. According to the O’Hora and Maglieri (2006) model, performance will increase until goal levels have been reached, and then will cease. In the current study performance levels increased with the introduction of the low goal (see Figure 3). Interestingly, scores for those participants that performed above 20 during baseline continued to increase following the introduction of the low goal condition. It was anticipated that performance would either remain stable or reduce to the goal level, if baseline performance had been above 20 correct entries. Of those participants that scored above 20 in baseline, 44% continued a positive trend through the low goal condition, and 71% demonstrated a level increase with the introduction of the low goal condition. A possible reason for the continuing rise in performance is that participants were learning the task through baseline, and continued a learning trend into the low goal condition. The effect of such “skill acquisition” was not measured in the paradigm, and future studies might seek to implement an extended baseline,
in addition to longer work sessions. With an extended baseline it may be possible to establish a stable rate of responding before introducing the goal condition. This may ensure that any increase in variability of performance upon introduction of the goal is produced by presentation of the goal statement and not simply a learning trend.

The current study highlights the contribution that basic research can make to understanding how goal setting interventions work in industry. Short session length and limited exposure to work sessions are limitations of the current paradigm; however the results provide a basis for expanding future research in the area. The limitations highlighted above should drive future studies, particularly investigations of persistence or deterioration of goal-directed behavior over time when assigned an unattainable goal. Prior research has demonstrated the utility of goal setting interventions (e.g. Latham, 2003; Jackson & Zedeck, 1982); however the current study constitutes a first step in examining the behavioral processes behind how goal setting interventions work. By characterizing the dynamics of goal-directed behavior, we can begin to develop more sensitive interventions to encourage employee productivity.
References


Figure 1. Work task screen. The numbers in the figure illustrate the steps that the participant was required to take to complete the task.
Figure 2. End-of-condition screen showing a count of correctly entered patients’ data for each of the work periods in that condition. The screen also shows the amount of errors in the condition, along with the average correctly entered data per work period, and the total patients correctly entered for all of the conditions.
Figure 3. Average performance for both the high goal first group and the low goal first group.
Table 1

t-values obtained in planned comparison t-tests, comparing overall phase means of correct responses for the high goal first group

<table>
<thead>
<tr>
<th></th>
<th>High Goal</th>
<th>Second Baseline</th>
<th>Low Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Baseline</td>
<td>-6.69* (p &lt; 0.001)</td>
<td>-7.63* (p &lt; 0.001)</td>
<td>-6.16* (p &lt; 0.001)</td>
</tr>
<tr>
<td>High Goal</td>
<td></td>
<td>-0.69 (p = 0.51)</td>
<td>0.18 (p = 0.86)</td>
</tr>
<tr>
<td>Second Baseline</td>
<td></td>
<td></td>
<td>0.75 (p = 0.47)</td>
</tr>
</tbody>
</table>

Note. Hochberg’s (1988) procedure for multiple tests of significance was used to control for multiple comparisons (each for 12 degrees of freedom, critical p-value = 0.0125).

* p < 0.0125.
Table 2

_t-values obtained in planned comparison t-tests, comparing overall phase means of correct responses for the low goal first group_

<table>
<thead>
<tr>
<th></th>
<th>Low Goal</th>
<th>Second Baseline</th>
<th>High Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Baseline</td>
<td>-3.43* ( p = 0.005 )</td>
<td>-3.84* ( p = 0.002 )</td>
<td>-5.13* ( p &lt; 0.001 )</td>
</tr>
<tr>
<td>Low Goal</td>
<td>-0.43 ( p = 0.68 )</td>
<td>-2.54 ( p = 0.03 )</td>
<td></td>
</tr>
<tr>
<td>Second Baseline</td>
<td></td>
<td>-2.29 ( p = 0.04 )</td>
<td></td>
</tr>
</tbody>
</table>

_Note._ Hochberg’s (1988) procedure for multiple tests of significance was used to control for multiple comparisons (each for 12 degrees of freedom, critical \( p \)-value = 0.0125).

* \( p < 0.0125 \).
**Table 3**

*Overall mean percent correct scores (with standard deviation in brackets) for each phase, for the high goal first group and the low goal first group*

<table>
<thead>
<tr>
<th>% correct</th>
<th>Baseline 1</th>
<th>High Goal</th>
<th>Baseline 2</th>
<th>Low Goal</th>
<th>Choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline 1</td>
<td>94 (0.07)</td>
<td>94 (0.06)</td>
<td>94 (0.08)</td>
<td>92 (0.08)</td>
<td>92 (0.08)</td>
</tr>
<tr>
<td>Baseline 1</td>
<td>93 (0.05)</td>
<td>95 (0.06)</td>
<td>95 (0.04)</td>
<td>95 (0.05)</td>
<td>93 (0.03)</td>
</tr>
</tbody>
</table>
### Table 4
*Individual Means, Standard Deviation (SD), and Trend for each participant in the high goal first group*

<table>
<thead>
<tr>
<th></th>
<th>Baseline 1</th>
<th>High Goal</th>
<th>Baseline 2</th>
<th>Low Goal</th>
<th>Choice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Trend</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>1</td>
<td>15.00</td>
<td>2.00</td>
<td>↑ (1)</td>
<td>17.33</td>
<td>1.53</td>
</tr>
<tr>
<td>2</td>
<td>22.33</td>
<td>1.53</td>
<td>↑ (1.5)</td>
<td>28.67</td>
<td>0.58</td>
</tr>
<tr>
<td>3</td>
<td>20.00</td>
<td>0.00</td>
<td>→ (0)</td>
<td>26.33</td>
<td>0.58</td>
</tr>
<tr>
<td>4</td>
<td>26.67</td>
<td>1.15</td>
<td>↑ (1)</td>
<td>31.00</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>26.00</td>
<td>4.00</td>
<td>↑ (4)</td>
<td>31.00</td>
<td>2.65</td>
</tr>
<tr>
<td>6</td>
<td>15.67</td>
<td>0.58</td>
<td>→ (0)</td>
<td>19.00</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>25.33</td>
<td>0.58</td>
<td>↑ (0.5)</td>
<td>27.00</td>
<td>1.73</td>
</tr>
<tr>
<td>8</td>
<td>24.00</td>
<td>1.00</td>
<td>↑ (0.5)</td>
<td>25.33</td>
<td>0.58</td>
</tr>
<tr>
<td>9</td>
<td>15.33</td>
<td>7.02</td>
<td>↑ (3)</td>
<td>23.33</td>
<td>2.08</td>
</tr>
<tr>
<td>10</td>
<td>21.33</td>
<td>0.58</td>
<td>→ (0)</td>
<td>22.33</td>
<td>1.53</td>
</tr>
<tr>
<td>11</td>
<td>15.67</td>
<td>2.08</td>
<td>↑ (0.5)</td>
<td>22.33</td>
<td>2.89</td>
</tr>
<tr>
<td>12</td>
<td>21.33</td>
<td>3.06</td>
<td>↑ (3)</td>
<td>25.67</td>
<td>0.58</td>
</tr>
<tr>
<td>13</td>
<td>23.67</td>
<td>4.16</td>
<td>↑ (4)</td>
<td>27.00</td>
<td>2.65</td>
</tr>
</tbody>
</table>

*Note.* Each row denotes a participant, with 13 participants in the group. A participant row gives mean, SD and slope value for each condition, as well as the chosen goal condition in the choice section (LG is low goal and HG is high goal). For example, Participant 1 scored a mean of 15 ($SD = 2$) in baseline, and demonstrated an increasing trend throughout baseline.
<table>
<thead>
<tr>
<th>Participant</th>
<th>Baseline 1 Mean</th>
<th>Baseline 1 SD</th>
<th>Baseline 1 Trend</th>
<th>Low Goal Mean</th>
<th>Low Goal SD</th>
<th>Low Goal Trend</th>
<th>Baseline 2 Mean</th>
<th>Baseline 2 SD</th>
<th>Baseline 2 Trend</th>
<th>High Goal Mean</th>
<th>High Goal SD</th>
<th>High Goal Trend</th>
<th>Choice Mean</th>
<th>Choice SD</th>
<th>Choice Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29.67</td>
<td>2.08</td>
<td>↑ (1.5)</td>
<td>33.00</td>
<td>1.73</td>
<td>↑ (1.5)</td>
<td>31.00</td>
<td>1.00</td>
<td>↓ (-0.5)</td>
<td>32.67</td>
<td>2.31</td>
<td>→ (0)</td>
<td>LG</td>
<td>32.67</td>
<td>0.58</td>
</tr>
<tr>
<td>2</td>
<td>18.33</td>
<td>1.15</td>
<td>↓ (-1)</td>
<td>19.67</td>
<td>2.08</td>
<td>↓ (-2)</td>
<td>18.33</td>
<td>1.53</td>
<td>↓ (-1)</td>
<td>20.67</td>
<td>2.08</td>
<td>↓ (-2)</td>
<td>HG</td>
<td>20.33</td>
<td>1.53</td>
</tr>
<tr>
<td>3</td>
<td>19.33</td>
<td>2.52</td>
<td>↑ (1.5)</td>
<td>22.67</td>
<td>0.58</td>
<td>↑ (0.5)</td>
<td>21.67</td>
<td>1.53</td>
<td>↓ (-1)</td>
<td>23.67</td>
<td>2.08</td>
<td>↓ (-1.5)</td>
<td>HG</td>
<td>23.67</td>
<td>1.15</td>
</tr>
<tr>
<td>4</td>
<td>20.00</td>
<td>0.00</td>
<td>→ (0)</td>
<td>22.00</td>
<td>0.00</td>
<td>→ (0)</td>
<td>21.67</td>
<td>1.53</td>
<td>↓ (-1.5)</td>
<td>23.67</td>
<td>1.53</td>
<td>↓ (-1.5)</td>
<td>LG</td>
<td>22.67</td>
<td>1.53</td>
</tr>
<tr>
<td>5</td>
<td>20.00</td>
<td>2.65</td>
<td>↑ (1)</td>
<td>25.00</td>
<td>1.00</td>
<td>↓ (-0.5)</td>
<td>27.33</td>
<td>1.15</td>
<td>↑ (1)</td>
<td>24.00</td>
<td>2.65</td>
<td>↓ (-0.5)</td>
<td>LG</td>
<td>28.33</td>
<td>3.51</td>
</tr>
<tr>
<td>6</td>
<td>18.00</td>
<td>1.00</td>
<td>↑ (0.5)</td>
<td>20.67</td>
<td>1.53</td>
<td>↑ (1.5)</td>
<td>21.67</td>
<td>0.58</td>
<td>↓ (-0.5)</td>
<td>20.67</td>
<td>1.53</td>
<td>↑ (1.5)</td>
<td>LG</td>
<td>19.33</td>
<td>1.53</td>
</tr>
<tr>
<td>7</td>
<td>21.67</td>
<td>3.51</td>
<td>↑ (3.5)</td>
<td>26.33</td>
<td>1.15</td>
<td>→ (0)</td>
<td>25.33</td>
<td>1.53</td>
<td>↓ (-1.5)</td>
<td>33.67</td>
<td>1.15</td>
<td>↓ (-1)</td>
<td>HG</td>
<td>29.00</td>
<td>2.65</td>
</tr>
<tr>
<td>8</td>
<td>22.33</td>
<td>1.53</td>
<td>↑ (1)</td>
<td>23.33</td>
<td>2.52</td>
<td>↑ (1.5)</td>
<td>21.33</td>
<td>2.52</td>
<td>↑ (2.5)</td>
<td>23.00</td>
<td>4.36</td>
<td>↑ (4)</td>
<td>HG</td>
<td>21.67</td>
<td>7.23</td>
</tr>
<tr>
<td>9</td>
<td>18.00</td>
<td>1.00</td>
<td>↓ (-1)</td>
<td>19.67</td>
<td>1.15</td>
<td>→ (0)</td>
<td>19.33</td>
<td>0.58</td>
<td>→ (0)</td>
<td>22.33</td>
<td>0.58</td>
<td>↑ (0.5)</td>
<td>LG</td>
<td>20.33</td>
<td>0.58</td>
</tr>
<tr>
<td>10</td>
<td>18.33</td>
<td>2.08</td>
<td>↑ (2)</td>
<td>21.00</td>
<td>0.00</td>
<td>→ (0)</td>
<td>22.67</td>
<td>2.08</td>
<td>↑ (1.5)</td>
<td>24.00</td>
<td>1.00</td>
<td>↓ (-1)</td>
<td>LG</td>
<td>23.67</td>
<td>2.08</td>
</tr>
<tr>
<td>11</td>
<td>22.67</td>
<td>2.52</td>
<td>↑ (1)</td>
<td>26.33</td>
<td>1.15</td>
<td>↑ (1)</td>
<td>28.00</td>
<td>2.65</td>
<td>↑ (2)</td>
<td>29.33</td>
<td>0.58</td>
<td>→ (0)</td>
<td>LG</td>
<td>26.67</td>
<td>2.52</td>
</tr>
<tr>
<td>12</td>
<td>31.33</td>
<td>2.52</td>
<td>↓ (-1.5)</td>
<td>27.00</td>
<td>2.65</td>
<td>↓ (-0.5)</td>
<td>32.67</td>
<td>2.08</td>
<td>↑ (2)</td>
<td>34.00</td>
<td>0.00</td>
<td>→ (0)</td>
<td>HG</td>
<td>34.00</td>
<td>1.00</td>
</tr>
<tr>
<td>13</td>
<td>17.00</td>
<td>1.00</td>
<td>↓ (-0.5)</td>
<td>18.67</td>
<td>0.58</td>
<td>↓ (-0.5)</td>
<td>17.67</td>
<td>0.58</td>
<td>↓ (-0.5)</td>
<td>18.33</td>
<td>3.06</td>
<td>↑ (3)</td>
<td>LG</td>
<td>19.00</td>
<td>2.00</td>
</tr>
</tbody>
</table>

Note. Each row denotes a participant, with 13 participants in the group. A participant row gives mean, SD and slope value for each condition, as well as the chosen goal condition in the choice section (LG is low goal and HG is high goal).
Figure 4. Scores from one participant, highlighting a decrease in performance during return to baseline condition.
Figure 5. Scores from one participant, highlighting a decrease in performance within the high goal condition.
Figure 6. Scores from one participant, highlighting a performance level increase with the introduction of the high goal.
APPENDIX

During this experiment you will be coding data for a medical company. You will examine the data provided for each patient and decide whether it is within range or out of range for each patient. To do this, you will first have to input the patient’s ID number (in the blue box to your left) in the PATIENT ID box in the center of the screen. Next, check the patient’s gender, and his or her QT interval number and compare it to the relevant box in the center of the screen (i.e. for female patients, check the FEMALE range, for males, the MALE range). If the patient’s QT interval is within the range for his or her gender, then click the dot next to ‘within range’. If not, click the dot next to ‘out of range’. When you are satisfied that you have made the right decision, click on SUBMIT to process this patient’s file. The experiment is divided into a number of sessions. Sometimes you will be working unsupervised. Other times, a virtual manager will ask you to complete a certain number of patients correctly. Well, that concludes the instructions for this experiment. The experimenter will deal with any further queries you have at this point. Remember to let the experimenter know if you need a break for any reason. In this first session there will be no virtual manager present. If you are ready to start the experiment, just click OK and then START NEXT SESSION. Best of luck!