Determining the Importance of Self-Evaluation on the Goal-Performance Effect in Goal Setting: Primary Findings

Bruce Martin* School of Business, University College Dublin, Ireland
Jeffrey McNally Faculty of Business Administration, University of New Brunswick, Canada
Simon Taggar School of Business and Economics, Wilfrid Laurier University, Canada

Acknowledgements: We gratefully acknowledge the financial support of the Centre for Strategic Leadership, School of Business and Economics, Wilfrid Laurier University, and thank Pak Hong Wong and Sarah Vermunt for their help in collecting and coding data.

* Corresponding author
Determining the importance of self-evaluation on the goal-performance effect in goal setting: Primary findings

ABSTRACT

Although goal-setting theory is among the most studied theories in organizational behavior and work motivation, the underlying motivations that drive the goal-performance effect have received less attention. We examined the role of self-evaluation in generating the goal-performance effect via blind testing in a laboratory experiment, in which participants ($N = 405$) performed an idea generation task under conditions eliminating the potential for external-evaluation. Designed to replicate and extend the work of Harkins, White, and Utman (2000), our results indicate that self-evaluation plays a role in generating the goal-performance effect, and that the pursuit of self-knowledge as well as self-validation plays a role in motivating self-evaluation. These findings support hypotheses that are consistent with goal setting theory and social comparison theory, and are contrary to Harkins et al. (2000). Implications for theory boundaries and work motivation are discussed.

**Keywords:** goal setting theory; self-evaluation; self-knowledge; social comparison; external evaluation; work motivation
INTRODUCTION

Goal-setting theory (Locke & Latham, 1990) is one of the most widely tested theories in the field of work motivation and organizational behaviour more generally, with more than 1,000 studies supporting its development over the past forty years (Mitchell & Daniels, 2003). According to Locke and Latham (1990; 2002), the underlying premise of goal-setting theory is that (a) goals function as immediate determinants of behaviour, (b) specific and difficult goals lead to high levels of absolute performance, and (c) for this to happen individuals must be motivated by a perceived discrepancy between current performance and a desired goal. Self-evaluation occurs when individuals compare their desired goal state with current performance levels (Kanfer, 1990). One problem with not having a clear goal (e.g., adopting a do-your-best goal) is the ambiguity as to what constitutes effectiveness. It is difficult to self-evaluate where there is no clear goal.

Until recently, it was generally accepted that people are motivated by the opportunity for self-evaluation via comparison of their performance with a specific difficult criterion (Locke & Latham, 1990). This view of self-evaluation as a fundamental motivating influence stems from Festinger’s (1954) social comparison theory, which suggests that people are driven to self-evaluate to avoid incorrect beliefs, or more generally to seek self-knowledge. According to Festinger (1954), self-evaluation occurs via objective/non-social means, if such means are available, or via social comparison, if objective/non-social means are not available. In other words, people have a basic desire to evaluate their abilities relative to some standard to learn about themselves (i.e., self-knowledge). The ideal standard is objective, such as being able to jump four feet, versus subjective, such as being able to jump quite far. Further, in Festinger’s view, the ideal objective standard is one found in the physical world, without reference to other
humans (i.e., objective/non-social), but if such standards are not available, people will use the abilities of others as a standard (i.e., social comparison).

Goethals and Darley (1987) modified Festinger’s social comparison theory, by demonstrating that the need for self-validation is also at play, and is likely equal to the need for self-knowledge. In this view, people may be motivated to self-evaluate in order to gain accurate information about their ability (self-knowledge), and/or to confirm that they are as good as their peers (self-validation). Given Goethals and Darley’s modification of Festinger’s theory, an investigation of the role of self-evaluation in goal setting must also investigate whether the underlying motivation for self-evaluation is generated by the need for self-knowledge versus self-validation.

The goal-performance effect has been demonstrated in a wide variety of contexts: both field and laboratory settings; over time spans varying from one minute to 25 years (Locke & Latham, 2006; Latham & Baldes, 1975; Howard & Bray, 1988); using individuals, groups and organizations as the level of analysis (Locke & Latham, 2002; Weldon, Jehn, & Pradhem, 1991); with assigned, self-set, and participatively-set goals (Latham, Erez, & Locke, 1988; Latham & Frayne, 1989); in studies across Asia, Australia, Europe, and North America (Locke & Latham, 2002). As an inductive program of research, its findings are widely supportive of the goal-performance effect in a variety of contexts and across time (Locke & Latham, 2013a).

Self-evaluations influence self-regulation, with individuals monitoring their behavior, comparing their actions with meaningful goals and adjusting their behavior to minimize discrepancies between present actions and desired outcomes (Cellar, 2011; Carver & Scheier, 1998). Goal setting is primarily a goal producing discrepancy process and secondarily a discrepancy reduction process (Locke & Latham, 2013a). The discrepancy between current performance and the desired goal can impact performance by focusing attention and effort
(Rothkopf & Billington, 1979), increasing energy applied to a task (Bandura & Cervone, 1983), increasing persistence (Laporte & Nath, 1976), and increasing the identification and use of relevant knowledge and strategies (Wood & Locke, 1990). For instance, an individual who wants to generate 50 ideas, and believes (s)he is unlikely to achieve this goal unless (s)he works hard, may (a) pay greater attention and give more effort to generating ideas and away from activities that do not generate ideas, (b) increase effort because (s)he anticipates the self-satisfaction of achieving the goal, (c) continue to work rather than stopping because the goal is not yet achieved, and (d) consider different idea generation strategies to achieve the goal. The goal-performance relationship is moderated by four main factors: goal commitment, feedback regarding progress in achieving goals, task complexity, and situational resources (Locke & Latham, 2002; 2013a). In the present experiment, we focused on a facet of goal commitment.

Self-evaluation influences goal commitment. Goal commitment is an individual’s determination or motivation to invest effort to achieve a desired goal (Latham & Locke, 1990, 2006; Fishbach, Zhang, & Koo, 2009; Koo & Fishbach, 2008), and is influenced by the perceived importance of the goal (i.e., goal value) and the perceived likelihood of attaining the desired goal (i.e., goal expectancy; Monzani, et al., 2015).

Although “a great deal is known about the consequences of goal commitment, far less is known about its antecedents” (Burkley, Anderson, Curtis, & Burkley, 2013; p. 667). Moreover, the specific role of self-evaluation in generating the goal-performance effect is still somewhat controversial. Locke and Latham (1990; 2013b) assert that self-evaluation plays a role in generating the goal-performance effect, and studies such as Zimmerman and Campillo (2003) have shown that self-evaluations significantly influence the forethought phase of the goal-performance relationship, during which individuals are expected to set their performance goals – and their level of difficulty – on their own. Yet Harkins, White, and Utman (2000) call into
question the role of self-evaluation in generating the goal-performance effect, and goal setting scholars have called for more work designed to understand the role of self-evaluation in goal setting more generally (Day & Unsworth, 2013).

Harkins et al., (2000) point out that Locke and Latham’s (1990) claim that the goal-performance effect is motivated by self-evaluation has little, if any, empirical support, as nearly all goal-setting studies included means for both internal evaluation (self-evaluation) and external evaluation (evaluation by others, such as an experimenter). Harkins and colleagues (Harkins et al., 2000; White, Kjelgaard, & Harkins, 1995; Szymanski & Harkins, 1987) examined the impact of self-evaluation versus external evaluation in generating the goal-performance effect. Their findings purport to show that “the potential for self-evaluation does not contribute at all” to the goal-performance effect (p. 103).

The present study was designed to examine whether Harkins et al.’s (2000) findings should be interpreted as showing that the potential for self-evaluation does not contribute at all to the goal-performance effect, as the authors claim. We believe this is important because we are answering their unchallenged finding that self-evaluation plays little to no role in the goal-performance relationship. This conflicts with scholarly claims that self-evaluation is critical to the goal-performance effect (e.g., Ashford & De Stobbeleir, 2013). As such, we propose an experiment to examine whether 1) self-evaluation plays a role in generating the goal-performance effect, as Locke and Latham (1990) claimed and Harkins et al. (2000) suggest they have refuted, and 2) whether self-knowledge plays a role in motivating self-evaluation, as Goethals and Darley (1987) claimed and Harkins et al. (2000) suggest they have refuted. A laboratory experiment, in which participants (N = 405) performed an idea generation task, was undertaken to first attempt to replicate the results of Harkins et al. (2000) and then further examine the impact of self-evaluation and self-knowledge on the goal-performance effect.
EXAMINATION OF HARKINS ET AL. (2000) FINDINGS

Harkins et al. (2000) report findings from three experiments examining the performance of undergraduate university students on an idea generation task involving generating uses for a knife. Their third experiment involved three manipulations, designed to isolate conditions for 1) goal setting, 2) external evaluation versus self-evaluation of performance, and 3) self-knowledge versus self-validation within the self-evaluation condition. The first manipulation varied experimenter demands by urging participants to strive for a goal of generating 40 uses for a knife versus saying nothing about meeting a goal, but simply encouraging participants to do their best (goal-setting/do-your-best). The second manipulation involved leading half the participants to believe that the experimenter would evaluate their specific performance, while the other half were led to believe that no one, including the experimenter, would know their specific performance (external evaluation/self-evaluation). The third manipulation varied the criterion for self-evaluation, by justifying the goal to be achieved as that which had been attained by similar participants versus dissimilar participants (self-validation/self-knowledge).

Harkins et al. (2000) used the external evaluation/do-your-best condition scores as the baseline for comparing all other outcomes. They considered only those outcomes found to be significantly above the performance level achieved in this condition as (potential) indicators of a goal-performance effect. Given this strategy for interpreting their results, only the two external evaluation/goal setting conditions showed a goal-performance effect. The highest scores in all other conditions equaled (statistically) the baseline do-your-best condition. This led them to conclude that 1) external evaluation is an important mediator of the goal-performance effect, 2) self-evaluation does not mediate the goal-performance effect, 3) self-validation may play a role in
improving task performance, but not enough to create the goal-performance effect, and 4) a result and implied conclusion that self-knowledge does not play a role in improving task performance.

We now outline our research, based on an evaluation of Harkins et al.’s (2000) conclusions and an analysis of extant literature as it relates to these conclusions. Their first conclusion was that external evaluation is an important mediator of the goal-performance effect. One problem with this conclusion is that significant performance improvements were demonstrated by those externally evaluated, versus those not externally evaluated in all conditions, including do-your-best conditions. That is, even when no goals were set, external evaluation appeared to generate significant performance increases in their study. This is consistent with extant literature (Locke & Latham, 1990; Latham & Lee, 1986; Likert 1967; French & Raven, 1959), which shows that compliance with legitimate authority or power often plays an important role in increasing task performance. According to Locke and Latham (1990):

Legitimate authority certainly exists in both laboratory and field settings. It may account for the high degree of generalization of results found between the two settings (Latham & Lee, 1986). In the laboratory, the experimenter is typically seen as an authority figure. An experiment, by its nature, is a “demand situation” (Orne, 1962). Subjects, who are almost always volunteers (and may get extra course credit), come to the experimental situation with the mental set: “Here I am; tell me what to do.” (p. 133)

Given this, one may interpret Harkins et al.’s (2000) results as confirming extant literature (e.g., Lock & Latham, 1990) by showing that external evaluation can have a significant motivational impact on task performance, regardless of the presence of goal-setting conditions. In other words, external evaluation plays a role in improving performance, but does not by itself generate the goal-performance effect. Accordingly, we hypothesize:
Hypothesis 1 (H1): Participants subject to experimenter evaluation will outperform those not subject to experimenter evaluation, regardless of goal assignment.

Harkins et al.’s (2000) second conclusion was that self-evaluation does not mediate the goal-performance effect. This conclusion was derived from comparing the performance in goal-setting conditions without external evaluation to performance in do-your-best conditions with external evaluation. Given our discussion above about the potential motivating role of external evaluation on increasing performance, this conclusion needs further examination. Indeed, it may be derived from a comparison that does not actually measure performance related to goal setting. We argue that the goal setting/self-evaluation conditions can only be compared with the do-your-best/self-evaluation conditions and not with the do-your-best/external evaluation conditions. Thus we suggest Harkins et al.’s finding in this area has no bearing on goal-setting, but rather suggests, in line with social loafing learning (e.g., Karau & Williams, 1993; Williams & Karau, 1991), that external evaluation can be a powerful motivator, even without goal-setting conditions. Thus, consistent with goal-setting theory, which proposes that specific, difficult goals lead to higher task performance, we hypothesize:

H2: Among those participants subject to no-experimenter-evaluation, those who receive specific difficult goal instructions based on a similar comparison group will outperform those who receive only a do-your-best instruction.

Lastly, we build upon Harkins et al.’s (2000) third conclusion that self-evaluation via social comparison (self-validation, i.e., self-evaluation based on peer performance) plays a role in improving task performance, but not enough to create the goal-performance effect. We do this to examine their implied fourth conclusion that self-evaluation without social comparison (self-knowledge, i.e. self-evaluation based on objective, non-social means) does not play a role in
improving task performance, and thus also does not play a role in creating the goal-performance effect. Our concern is that the specific operationalization of tasks in Harkins et al. (2000) may underestimate the role of self-evaluation in generating the goal-performance effect. It is possible that the experimental conditions established by Harkins et al. (2000) effectively eliminated any basis for generating task motivation or goal commitment, other than external evaluation by the experimenter and social comparison (self-validation). As such, results showing that task performance in self-knowledge conditions are no better than do-your-best levels, may not be surprising given the experimental design.

Consider the situation in which the participants in the self-evaluation condition that relies on self-knowledge would find themselves. They are undergraduate students who have agreed to participate in a study to gain course credit. The task assigned to them involves generating uses for a knife, something that is unlikely to be inherently engaging to most people, as noted by Harkins et al. (2000) and Szymanski and Harkins (1987). Further, they have been told that 1) the quality of their responses does not matter, only the quantity is of concern; 2) no one, including the experimenter, will know how they perform; and 3) the only performance comparison available is that of third-year PhD students, people not just dissimilar, but likely to be judged much more capable than themselves (Harkins et al., 2000). In such a situation, participants may lack the motivation to perform well on a task, and thus it may not be surprising to find that task performance levels were similar to do-your-best levels. Locke and Latham (1990, p. 124) state “it is virtually axiomatic that a goal that a person is not really trying for is not really a goal and therefore cannot have much effect on subsequent action”.

Accordingly, we suggest that this condition is not measuring the effect that self-knowledge has on generating a goal-performance effect, but rather the ability to generate the goal-performance effect when there is little task motivation. Hence, we suggest that these
findings may not provide an appropriate basis for drawing conclusions with regard to the motivating potential of self-knowledge, nor the role that this might play in generating the goal-performance effect.

In order to examine the self-knowledge condition it is necessary to ensure that participants are given at least some means for generating task motivation, such as a stimulus that makes the task outcome appear to have some value (value-of-outcome-specified). Of course, such a unique stimulus would need to be applied equally to all comparison conditions. Consistent with this, Harkins et al. (2000) do acknowledge that presenting participants with evidence of the value of the idea generation task may overcome concerns with the lack of inherent interest the task holds and the lack of justification provided to participants for why the task should be performed. Addressing this would better recognize Locke and Latham’s (2002) assertion that goal attainment must hold importance to people, either because they have set the goal themselves or because they have been given a good reason or rationale for why the goal should be attained.

Utman and Harkins (2010) attempted to address the issue of goal attainment importance in a replication and extension study building on Harkins et al. (2000), by introducing a prime that tied performance on the knife-uses idea generation task to intellectual ability, telling participants that past research showed that the more knife-uses people generated the higher their academic ability. Their results indicated that with this prime, those in self-validation conditions achieved the goal setting effect.

Proceeding on the basis that a stimulus—specifying the value of the task outcome—can be introduced, we continue to develop hypotheses for testing both self-validation and self-knowledge in goal-setting contexts. Consistent with Goethals and Darley’s (1987) version of social comparison theory, proposing that people may be motivated to self-evaluate in order to
confirm that they are as good as their peers (self-validation); and with Locke and Latham’s (1990) assertion that self-evaluation is a motivating factor in goal-setting, we hypothesize:

**H3: Among those participants subject to value-of-outcome-specified conditions and no-experimenter-evaluation, those who receive specific difficult goal instructions based on a similar comparison group will outperform those who receive only a do-your-best instruction.**

Unlike Harkins et al. (2000), Utman and Harkins (2010) examined only self-validation as a means for generating the desire to self-evaluate, and not self-knowledge, perhaps because they viewed the Harkins et al. (2000) results as having proven that self-knowledge does not generate the desire to self-evaluate. In the current study, we examine the role of self-knowledge as well as self-validation in generating the desire for self-evaluation. This is in line with Festinger’s (1954) social comparison theory, which suggests that people are driven to self-evaluate in order to seek self-knowledge. Goethals and Darley’s (1987) emphasize the addition of self-validation as a motivator of self-evaluation within the context of Festinger’s social comparison theory, but they acknowledge that self-knowledge is also a motivating factor, and may be used when social comparisons are not possible. Although they do not delineate the underlying motivators of self-evaluation, between self-validation and self-knowledge, Locke and Latham (1990; 2013b) assert that self-evaluation is a motivating factor in goal-setting, operating through goal commitment. Thus, there is reason to expect that self-knowledge may play a role in motivating the goal-performance effect, but no study to our knowledge has properly assessed this phenomenon. Accordingly, we hypothesize:

**H4: Among those participants subject to value-of-outcome-specified conditions and no-experimenter-evaluation, those who receive specific difficult goal instructions**
based on a dissimilar group comparison will outperform those who receive only a do-your-best instruction.

METHOD

Participants

Participants were 405 undergraduate business students enrolled in Organizational Behaviour or Human Resources Management courses in a mid-sized Canadian university, who received course credit in exchange for their participation in the study. On average, participants were 20 years old (SD = 1.23) and 53% were male. Participants were randomly assigned to one of 12 conditions.

Procedure

In order to ensure a direct replication of Harkins et al. (2000), we first designed our study using their published procedures and then contacted the lead author to confirm that our interpretation was accurate (Harkins, 2009a). The study uses an idea generation task in order to examine performance on a task. The use of idea generation exercises, such as creating uses for a knife, is common in task performance studies of this sort (e.g., White et al., 1995; Szimanski & Harkins, 1987) as they are exercises that can be performed by a wide variety of participants without the need for long explanation or considerable prior skills or practice. Also, there is limited range restriction, allowing for considerable variance, which is important to examining work performance outcomes.

Our design follows Harkins et al. (2000) in all respects except, when a goal was given, we used a goal of 50 uses for a knife, rather than 40 uses for a knife as in Harkins et al. (2000). Our need for a higher goal was determined in a pilot study we conducted, which showed that a goal of 40 was inappropriate for our participants. Locke and Latham (1990) suggest that goals be set at a level that is attainable by approximately 10% of participants in laboratory experiments, in order
to fit with the difficult but attainable parameter. In our pilot study, we found that 54% of participants in the do-your-best condition generated 40 or more ideas, and that 12% generated 50 or more. This suggested that a goal of 40 uses for a knife was too low, and that 50 uses for a knife would be more appropriate. Thus, although our pilot replication study was accurate in all ways including sample size per cell, it did not create a valid representation of goal setting for this population, as the goal was set too low to motivate significant increases in performance levels.

The experiment was conducted by a male Masters student in his 20s and a female PhD student in her 30s, neither of whom was informed of the study’s hypotheses. Data were collected from one to six participants at a time. Participants were seated at individual workstations in a laboratory with walls that blocked their view of other participants. They also wore sound dampening headphones to minimize awareness of fellow participants’ writing activity. After being seated, all participants were asked to turn off cell phones and other electronic items. The experimenter explained that they would be asked to perform a number of idea generation task exercises. They completed a warm-up session in order to become familiar with an idea generation task. This warm-up session involved generating uses for a box over a three-minute time period. Before beginning the warm-up, all participants were told that the quality of uses they generate does not matter, only the quantity will be measured. Half of the participants were told that the experimenter would count the number of uses each one of them generates at the end of the session and then add their list of uses to a box containing the lists of uses generated by previous participants (external-evaluation group). The other half of the group was told to add their list of uses to a box containing the lists of uses generated by previous participants, before the experimenter enters the room (self-evaluation group). They were told that, as this box already contained other responses, the experimenter would have no way of knowing any one individual’s response.
After completing the warm-up exercise, participants were told that the main exercise involved generating uses for a knife over a 12-minute time period. At this point, the external evaluation/self-evaluation instructions were repeated. A third of participants (do-your-best group) were instructed “to come up with as many ideas for using a knife as you possibly can” and encouraged to “do your best.” Another third of participants (undergraduate-goal group) were told:

This experiment was run before with undergraduate students in Organizational Behaviour and Human Resources Management courses. They could generate up to 50 uses for a knife in 12 minutes. I would like you to use this as a goal for your own performance. Therefore, I would like you to work hard to try to reach 50 uses.

The remaining third of participants (graduate-goal group) were given the same 50-goal message as the undergraduate-goal group, but with “third-year Ph.D. students” replacing “undergraduate students in Organizational Behaviour or Human Resources Management courses” as the type of participants from which the goal was derived.

As an extension of Harkins et al. (2000), participants were further divided in half based on the value-of-outcome manipulation. One half of the participants (value-of-outcome-specified group) were told that their performance on this task is important because “a $500 donation will be made to the local United Way if the average results in this study are at least equal to previous studies.”

A charitable donation was chosen as the means for ensuring that those in the (no-experimenter-evaluation, dissimilar-comparison-goal group) condition who, as we described in some detail in the literature review, face the unusual situation of being asked to do a task that is inherently uninteresting, for which no one including the experimenter will know their performance, and about which they are given a goal based on a group that is dissimilar to them,
will still have some reason to view performing well on the task as meaningful. We avoided using a direct extrinsic reward, recognizing that such incentives have been shown to differentially impact based on individual differences, such as personality characteristics, for instance (e.g., Locke & Braver, 2008), and thus may impact results within groups.

The United Way was chosen as the donation recipient because it supports a broad array of social services, and a small informal survey of our target sample indicated that it was known to and would appeal to a broad selection of participants. The remaining half of the participants (value-of-outcome-not-specified group) were not told anything about the value of the outcomes they produce, thus following Harkins et al. (2000) procedure. At this point, the participants filled out a questionnaire, which included manipulation checks and demographic information.

After delivering the instructions to participants, the experimenter left the room and closed the door. After 12 minutes, the experimenter knocked on the door and either entered and counted the participants’ uses (external-evaluation group) or waited until the participants put their sheets into the box and then opened the door (self-evaluation group). At this point, the participants filled out a questionnaire, which included manipulation checks and demographic information. Manipulation checks included the following questions: “Did the experimenter examine the results of your individual performance on the ideation task? (yes/no)”; “Will the experimenter(s) be able to determine: 1. Your specific results on the ideation task? 2. Only the average results of all participants? (Please choose only one answer)”; “Were you told that a donation would be made based on the level of task performance of all participants in this study? (yes/no)”; “If you were told that a donation will be made, to what organization will the donation be given? 1. United Way, 2. Habitat for Humanity, 3. Food Bank, 4. I Don't Remember, 5. Not Applicable (Please choose only one
answer); “Were you given a specific goal to achieve on this ideation task? (yes/no)”
“If you were given a goal, what was it? (Please write the number in digits)”
“If you were given a goal, what was the goal based on (why was that level chosen)?
1. It was based on the level that other undergraduate students achieved.
2. It was based on the level that other third-year Ph.D. students achieved.”

RESULTS
Manipulation check responses were compared with group assignment and those participants whose answers did not match the group to which they were assigned were removed from the sample, resulting in a reduction of 21 participants from the original 426 who joined the study. Hypotheses were tested using ANOVA and Tukey’s honestly significant difference test for pairwise comparisons.

To test H1, data were analyzed using a 2 (value-of-outcome-not-specified vs. value-of-outcome-specified) x 2 (external-evaluation vs. self-evaluation) x 3 (do-your-best vs. undergraduate-goal vs. graduate-goal) ANOVA, with a significant overall result, F (11, 394) = 4.14, p < .001. Results are reported in Table 1, indicating that only the goal factor (do-your-best vs. undergraduate-goal vs. graduate-goal) was significant, F (2, 403) = 18.65, p < .001, \(d = .431\). Thus, H1 was not supported.

-----------------------------
Insert Table 1 about here
-----------------------------

To test H2, we conducted a one-way ANOVA comparing do-your-best versus similar-comparison-goal versus dissimilar-comparison-goal for each condition specified. The test of
Hypothesis 2, which examines the goal variable outcomes within the value-of-outcome-not-specified/no-experimenter-evaluation condition, yielded a main effect, $F (2, 109) = 3.31 \ p < .05$, $d = .350$. However, as shown in Table 1, the similar-comparison-goal condition ($M = 41.2$), specified in H2 did not show a significant increase in knife-uses performance over the do-your-best goal condition ($M = 36.3$). Thus, H2 was not supported. It is interesting to note, however, that a significant increase in performance was found for the dissimilar-comparison-goal condition ($M = 43.7, \ p < .01$).

The test of H3 and H4, examining the goal variable outcomes within the value-of-outcome-specified/no-experimenter-evaluation condition, yielded a main effect, $F (2, 100) = 8.99 \ p < .001$, indicating support for the goal setting condition. As Table 1 shows, both H3 and H4 were supported, with increased performance over the do-your-best goal condition ($M = 35.4$) shown for both the similar-comparison-goal condition ($M = 46.9, \ p < .001$) and the dissimilar-comparison-goal condition ($M = 45.7, \ p < .01$).

**DISCUSSION**

**Theoretical Contributions**

This study was designed to answer two main questions: a) Does self-evaluation play a role in generating the goal-performance effect in the absence of potential for external evaluation; and b) does the pursuit of self-knowledge play a role in motivating self-evaluation? Our results provide support for the view that self-evaluation plays a motivating role in generating the goal-performance effect and, further, that the pursuit of self-knowledge plays a motivating role in generating the desire to self-evaluate. These findings provide a valuable contribution to theory by 1) providing support for Festinger’s (1954) social comparison theory, by showing that self-evaluation for the purpose of gaining self-knowledge does play a role in motivation, and 2)
providing support for Locke and Latham’s (1990) assertion that self-evaluation plays a motivating role in generating the goal-performance effect.

Results of this study differ from Harkins et al.’s (2000) findings in a number of important ways. We argue that these differences are consistent with the goal setting literature and result from methodological differences between the present study and Harkins et al. (2000). First, the goal-performance effect was exhibited by participants with and without experimenter evaluation, contrary to Harkins et al. (2000), but consistent with Locke and Latham (1990). This is an important finding in that it directly addresses the challenge that Harkins et al.’s findings made to goal setting theory; that their study was one of the few to examine the goal-performance effect in conditions without external evaluation, and the inability to obtain the goal-performance effect in such conditions suggested that goal setting results may be driven primarily by external evaluation influence. We argued that the specific experimental conditions established by Harkins et al. (2000) may have effectively eliminated any basis for generating participant task motivation in the absence of external evaluation by the experimenter and social comparison, because the task of identifying uses for a knife was unlikely to be inherently interesting to most people (i.e., low value). Second, contrary to Harkins et al., we show that the goal-performance effect occurs in both self-validation and self-knowledge conditions, which is consistent with theoretical expectations outlined initially by Goethals and Darley (1987) for performance motivation generally, and by Locke and Latham (1990) for goal commitment in goal setting theory specifically.

**Practical Implications**

Our findings may also contribute to management practice. Given that the application of goal-setting theory is widespread in the workplace (Locke & Latham, 2013b), furthering our
understanding of the underlying motivations that generate the goal-performance effect can better inform the specific workplace applications of goal-setting. Among the common methods used to build goal commitment in organizations is the use of authority figures to evaluate employee performance outcomes and attaching a monetary reward to goal achievement (Katzenbach & Khan, 2010). Although further research is needed before specific practitioner guidance can be offered, the present study suggests alternatives to both these prescriptions.

One example of the value of understanding how the goal-performance effect can be achieved without external evaluation is in workplace situations where providing external evaluation is costly or difficult. In such cases it may be advisable to lower monitoring levels or even eliminate them altogether. Indeed, recent scholarly work has begun to question the value of high levels of employee monitoring in organizations, with evidence that increases in employee monitoring can actually reduce effectiveness, as costs rise and cooperative signals from management are perceived by employees to decrease (Gnyawali & Madhavan, 2001). Employee monitoring also appears to be significantly associated with fear of job loss (Oz, Glass, & Behling, 1999), work dissatisfaction (Schleifer, Galinsky & Pan, 1995), and even chronic health disorders (Smith, Carayon, Sanders, Lim & LeGrande, 1992).

Over the past few decades there has also been growth in jobs that afford employees greater autonomy. For instance, in addition to jobs that have traditionally required autonomy from individuals (e.g., police officers and social workers working in the “field”), many organizations have reduced supervisory positions as part of becoming lean, designing work for teams, and decentralization so that supervisors have the ability to exercise greater authority over their employees (e.g., Kristensen, 2012; Mena Report, 2015; Sumathi, Kamalanabhan, & Thenmozhi, 2015). Under work conditions of reduced external evaluation, the goal-performance
effect may be aided in organizations through methods of gaining goal commitment other than external evaluation.

Specifically, we found that rather than do-your-best goals, individuals appear to benefit from specific goals with a valued outcome attached to them, along with self-validation and self-knowledge information. Appropriate self-validation information obtained via social comparison may encourage the belief that individuals are able to complete required actions (i.e., expectancy beliefs) and potentially motivate employees to avoid the stress and potential dissatisfaction if the set goal is not reached (i.e., value).

With respect to goal achievement value, for many organizations a lack of goal commitment is remedied by greater financial incentives for goal achievement. Money improves goal commitment, within certain parameters. For instance, a small monetary reward can undermine performance when perceived as insulting (Gneezy & Rustichini, 2000). Monetary rewards attached to goal achievement may impact the value facet of goal commitment but not the expectancy aspect of goal commitment. Interestingly, in the present study, a charitable donation was sufficient to make goal achievement motivating for participants when paired with the potential for self-validation or self-knowledge. With a low expectancy of success, individuals tend to disengage and abandon their goal (Carver & Scheier, 1998). Our results suggest that self-evaluation via either self-validation or self-knowledge can fuel motivation. With self-validation, when people feel that others with whom they might compare have achieved a specific goal they are likely to believe that they can also achieve the goal if they work hard. Similarly with self-knowledge if people feel that in pursuing a specific goal they can compare their performance with an objective, but meaningful performance criterion they are likely to be motivated to work hard. This is consistent with Lee (2015), who suggests that goals that are more difficult but still viewed as achievable actually promote greater goal commitment.
Many workplaces have removed social comparison information in the form of evaluations via rankings because they demoralize workers, foster backstabbing and favoritism, and foster power struggles among managers (Ovide & Feintzeig, 2013; Kwoh, 2012). Goals that are rewarded and provide potential for self-knowledge via objective information related to the task and the achievement needs, at the time of goal-setting, may be a way of providing information that improves performance without these detrimental effects.

The importance of self-evaluation in the goal-performance effect highlights the need for organizations to ensure the potential for self-evaluation is optimized. According to Garcia, Tor, and Schiff (2013) this can be achieved by increasing the relevance of the performance dimension during goal setting, the similarity of comparison competitors who are performing the task, and their relationship closeness to the individual, as well as the various individual differences variables relating to social comparison more generally. Situational factors impacting social comparison that can be manipulated by organizations include proximity to a standard (i.e., near the number 1 ranking vs. far away), the number of competitors (i.e., few versus many), and social category fault lines (i.e., disputes across versus within social categories).

Our results are consistent with Shalley, Oldham and Porac (1987) in suggesting that intrinsic motivation, arguably a precursor to performance in at least the “value of outcome not specified” condition, is highest when achievement goals are set regardless of external evaluation by an authority figure. That is, counter to Shalley and Oldham (1985), our findings support the notion that managers can help subordinates achieve relatively higher intrinsic motivation under relatively low personal control goal setting conditions (i.e., when goals are predetermined) compared to high personal control conditions (i.e., do your best goals). Intrinsic motivation is relevant to organizations because it is linked to individuals' creativity on tasks, risk taking, and product quality (Amabile, 1979).
Potential Limitations

We considered that our results related to external evaluation may have been impacted by the image of our experimenters. Most (83%) of the experiments reported here were conducted by a male Masters student in his 20s, while the remaining were conducted by a female PhD student in her 30s. It is possible that they did not appear to be authority figures. Although not conclusive, this would appear not to be an issue, relative to Harkins et al. (2000), as communication with the lead author indicates that a male graduate student in his 20s also conducted a significant portion of their study, although the lead author, a mature male, conducted the remaining portion (Harkins, 2009b).

We also considered that our results might be impacted by the fact that our experimenters were blind to the hypotheses of the study, whereas both experimenters used in Harkins et al. (2000) were aware of the study hypotheses (Harkins, 2009b). The impact of experimenter knowledge of study hypotheses has been shown to be considerable in the fields of medicine and psychology, for instance, but this issue has received only scant attention in work performance and organizational behaviour literature. It is difficult to assess the impact that experimenter knowledge of the study hypotheses might be having on outcomes of the current experiments, but future research should test the effects by controlling for this potentially important factor.

Another potential limitation relates to the generalizability of our results. By testing our hypotheses on only one specific population—undergraduate university students—we are unable to show that these phenomena apply widely in the population. Further, our use of laboratory experimental methods only, which allowed us to better control the variables to be examined, diminishes generalizability to real life settings. Future work should examine these phenomena amongst a broader sample of the population and in settings that better reflect real life contexts. Also, although self-efficacy and goal commitment are often measured in goal setting studies, we
do not include measures of these two variables, consistent with the study we replicated, Harkins et al. (2000). Future research incorporating measures of self-efficacy and goal commitment may provide further insight into the relationships examined here.

**Future Research**

The findings from this study also raise a number of interesting issues for further research. As well as the need to conduct replication studies to determine the robustness of the above findings, it is important to further examine the impact of external evaluation on task performance. Results from this study suggest that external evaluation may not always play an important role in task performance, which contradicts findings from Harkins et al. (2000). The fact that other extant literature (Locke & Latham, 1990; Latham & Lee, 1986; Likert 1967; French & Raven, 1959) indicates that external evaluation by an authority figure plays an important role in increasing task performance suggests the need for further study via experiments designed to specifically examine this phenomenon.

We also considered that our sample might be biased toward high achievement individuals. There is some indication that this might be the case, as evidenced by our need to increase the assigned goal from the level of 40 set by Harkins et al. (2000), due to the fact that so many participants in our sample, who were given no goal instructions, surpassed the 40 uses level in Experiment 1. It is also possible that the impact of external evaluation is moderated by the base level of self-motivation of participants or their intrinsic motivation in the task, such that at higher levels of self-motivation, the impact of external evaluation is diminished to the point of being insignificant. Further study in this area is warranted. In addition to base level self-motivation, it may also be valuable to examine self-efficacy and goal commitment levels among participants to better illuminate the phenomena under study and the impact of sampling differences stemming
from these characteristics. Future studies should also sample non-business students, and ideally a random sample of the general adult population to ensure that these findings are generalizable and to assess moderators related to such population differences.

**CONCLUDING REMARKS**

This paper provides insight into an understudied aspect of one of the most studied theories in work-motivation and organizational behaviour, goal-setting theory. We find evidence that supports Locke & Latham’s (1990) assertion that the desire to self-evaluate drives the higher relative performance outcomes of those striving to achieve specific, difficult goals versus those who strive to do their best. Further, we show that self-evaluation can be motivated by the need for either self-validation or self-knowledge, as predicted by Goethals and Darley’s (1987) extension of Festinger’s (1954) social comparison theory, but contrary to Harkins et al. (2000). We also demonstrate that this new understanding may open up new lines of investigation related to goal setting theory. For instance, by combining our learning with work that shows how the use of social comparisons can be increased, such as via self-activation (Stapel & Tesser, 2001) we can begin to examine ways to increase the likelihood that individuals set specific difficult goals for themselves and thus improve their self-directed task performance.

**REFERENCES**


### Table 1. Mean uses for a knife experiment

<table>
<thead>
<tr>
<th>Comparison</th>
<th>$M$</th>
<th>SD</th>
<th>$n$</th>
<th>$d$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Value of Outcome Not Specified</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do-your-best (base)</td>
<td>36.8</td>
<td>11.5</td>
<td>32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-validation goal</td>
<td>45.7</td>
<td>14.4</td>
<td>32</td>
<td>.68</td>
<td>[.17, 1.20]</td>
</tr>
<tr>
<td>Self-knowledge goal</td>
<td>41.1</td>
<td>8.3</td>
<td>32</td>
<td>.43</td>
<td>[-.08, .93]</td>
</tr>
<tr>
<td>Self-Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do-your-best (base)</td>
<td>36.3</td>
<td>10.5</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-validation goal</td>
<td>41.2</td>
<td>12.0</td>
<td>37</td>
<td>.44</td>
<td>[-.03, .90]</td>
</tr>
<tr>
<td>Self-knowledge goal</td>
<td>43.7</td>
<td>14.7</td>
<td>37</td>
<td>.58</td>
<td>[.11, 1.05]</td>
</tr>
<tr>
<td><strong>Value of Outcome Specified</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do-your-best (base)</td>
<td>36.7</td>
<td>9.9</td>
<td>32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-validation goal</td>
<td>45.9</td>
<td>11.6</td>
<td>32</td>
<td>.85</td>
<td>[.33, 1.37]</td>
</tr>
<tr>
<td>Self-knowledge goal</td>
<td>40.7</td>
<td>13.4</td>
<td>32</td>
<td>.34</td>
<td>[-.16, .84]</td>
</tr>
<tr>
<td>Self-Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do-your-best (base)</td>
<td>35.4</td>
<td>10.7</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-validation goal</td>
<td>46.9</td>
<td>14.5</td>
<td>34</td>
<td>.90</td>
<td>[.40, 1.41]</td>
</tr>
<tr>
<td>Self-knowledge goal</td>
<td>45.7</td>
<td>11.3</td>
<td>34</td>
<td>.94</td>
<td>[.43, 1.45]</td>
</tr>
</tbody>
</table>
